

TITLE Direct Submission
JOURNAL Submitted (28-MAY-1999) to the DDBJ/EMBL/GenBank databases. Hisao
Masaai, Institute of Medical Science, University of Tokyo,
Department of Developmental Biology; 4-6-1 Shirokanedai, Minato-ku,
Tokyo 108-8639, Japan (E-mail: hisao@ims.u-tokyo.ac.jp,
Tel: 81-3-5449-5661, Fax: 81-3-5449-5424)

FEATURES

source Location/Qualifiers

1..2719
/organism="Homo sapiens"
/db_xref="taxon:9606"
/cell_type="HeLa"
/chromosome="7"
/clone="H37"
/map="7q21"
518..1222
/gene="ASK/ H37 (hubb4)"
518..1222
/gene="ASK/ H37 (hubb4)"
/note="C-terminally truncated due to alternative splicing;
an alternative splicing variant of ASK, a regulatory
subunit for hucd7 kinase. This protein could bind to
hucd7 but may not be able to activate its kinase activity."
/codon_start=1
/product="activator of S phase kinase"
/protein_id="BAA78327.1"
/db_xref="GI:4996098"

polysite 932 a 522 c 607 g 658 t
BASE COUNT
ORIGIN

Query Match 100.0% Score 2719; DB 9; Length 2719;

Best Local Similarity 100.0% Pred. No. 0;
Matches 2719; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 aatcgagcagagctctctgagctcgaagacctaagcggcgagacggagagccgg 60
Db 1 AATTCGGCAGCGCTCTGTAGCTGCGCCAGACCTGAAGCGCGACGACGACCGCG 60
QY 61 gtctgagctgagagcaacgaaatgagagcgaggtagagcgagaaacacacactgca 120
Db 61 GTCTGAGACTGAGAGCAACGGAATGAGCGCGGTAGAGCGGAAACACACACTGCA 120
QY 121 ggcgaagcgaagcgcgagaaagcggcggtgagggcgaggcgcgacgagaa 180
Db 121 GCCCAAGCGAGCGCCGAGAGGACGCGCGGTGAGGCGCGCGCGCGCGGAA 180
QY 181 ggcgaagcgaagcgcgagcgagcgagcgagcgagcgagcgagcgagcgag 240
Db 181 GCGAGCGACGAGGCGCGAGCGCGGCGCGCGGTGAGCGGCGCGCGG 240
QY 241 cggcgctacggcgcgcgcgcgcgctgacgcttccaattccaacgcgcgagacca 300
Db 241 CCGCGGATTCGGCGCGCGCGCGGTGAGCGGCGCGGTTCGAATTCGCGCGGACCA 300
QY 301 ctgctgtgagcttgcgcttcctcctcgcgcgccttgagcgcgagatccggcccgaaac 360
Db 301 CTCGCTTGTGCTTGTGCGCTTCTCTCGCGCTTGAGCGCGGATTCGCGCGGAAAC 360
QY 361 ccgaactgcaagcgcgtaacttaactgctagaagcgtaactgagagagagagag 420
Db 361 CCGACCTGCAAGCGCGTACTTACTGCTAAGGCGCTACTGCGCGAAGGAGAGG 420
QY 421 cggcgctcctgtcaacagcgaggagcgagcgagcgagcgagcgagcgagcgag 480
Db 421 CGGCGCTCTGTCAACAGCGCGGAGCGCGTCTTCTGCGCGCTGCGCGGAGACT 480
QY 481 ttctcggagaccagcatgagtgccggcgagctgcattgaactccggagcattgagat 540

Db 481 TTCTCCGAGCCAGCATCTAGCTGCGCGGAGCTGCGCATGAATCTCGGAGCATGAGCAT 540
QY 541 ccacagtaagagacattccagagtggaatccagatccaaatgaaaaaagacagacac 600
Db 541 CCACAGTAAGAGCATCTTCCAGGCTGGAATCCAGCAATGAATGAAAAAAGACAGCATC 600
QY 601 lctgaactctctgaactgatcaagcgccagaaaaatccaaatgtaagccacttgggg 660
Db 601 TCTGAATCTCTGAAACTGATTAACAGCGCCAGAAAATCAATTAAGCCACTTGGGG 660
QY 661 aaaaatttctaccttgaacttcttctgtacacatattgaaaaactcaaaagacat 720
Db 661 AAAAGTATTTACTCTGACTTACCTTGTGACCATATCTGAAAACATTCAAAAGGACAT 720
QY 721 taagatcctggagcgagtgagaaatctccaagcaagatatacagttacttatttc 780
Db 721 TAAGATCTGGAGCGCGGCTTGAAGAAATTTCTACAGCAAGATATCACTTATTTTC 780
QY 781 aataagagagagcctaatttgcaacaacacttgagtgcaattctcctgtacaaagtc 840
Db 781 AATTAAGAGAGAGCTAAATTTGCCAAACCTTGGGCGAATTTCTCTGTACCAAGTCC 840
QY 841 agaatcgtatatactgagagaaacacttcaactcaccacacatgagtgagtgatc 900
Db 841 AGAATCTGATATCTGCGAGAAACACTTCACTCCATCCAGCATGATGAGATTCATT 900
QY 901 taagtcaacagacagtggtgttaagagagagaaatattatgattgaaaaagctatca 960
Db 901 TAACTTACCGACAGACAGTGTGTTTAAGCAGAGAAAATTATAGTTGAAAAAGCATCA 960
QY 961 ggaccatgatattatctctccaatagatatattcaaaagcctgtatagggagagaa 1020
Db 961 GGACCATGATTTTATCTCTCAATATGATATATCAATGCTTGTCACTGGGAGATTA 1020
QY 1021 aattctcatattgattgacattagatctacattgacaacaaaaaagtgattatc 1080
Db 1021 AATCTTCAATTTGATGACATTTAGATTAATGATTAATGAAACAAAAGAAATGATATTT 1080
QY 1081 actcaagaatactagacttctcagtagaagatgaggagcgaaagagtgtagtgatgaca 1140
Db 1081 ACTCAAGAAATCAAGTACTTCAATGAGATGGGGGCAAAAGAGTGTGATGGGCA 1140
QY 1141 aaaaacaaagacaggaagactcaaaaagccttctgtaaaggtgaaagatagagccaaag 1200
Db 1141 AAAAACAAAGAACGAAAGACTCAAAAAGCCTTTGTAAAGGTGGAAGATATGAGCAAG 1200
QY 1201 cctgtgagctcattgatgttagacaagcactatgatacgcaaaagcaactcaggttaa 1260
Db 1201 CCTGTGAGTGCATTTGTATGTAGCAAGCCATCTAGTATGCAAAAAGCAAACTCAGTTTAA 1260
QY 1261 ctaagaatccaaacagatgtagaagatgtagtgaacccaactcaactcagttgaa 1320
Db 1261 CTAAAGATCCAAACAGTATGCGCATTAATGTTGTGAACCTCAATTCACCTCAGTTGAA 1320
QY 1321 gagaagaagaaaaagatattgtgaatgttcttgcaagaataataaagataaagact 1380
Db 1321 GAGAGAGAGAAAGAAAGATATTTGTCAATTTCTTGCAGAAATATTAAGATCTAAGAACT 1380
QY 1381 caactcttaagttagcaacacagaaaacttgcacagagtaacagatataaagttgat 1440
Db 1381 CACCTTCTAAGTGAACACAGAACTTGTGACAGAGATACCAAGTATCAATTTGTAT 1440
QY 1441 gatattgatalaagtagtatttctgacttctgtgaatatgaaaagacacacactaaag 1500
Db 1441 GATATTGTATTAAGTATGTTTGTGACTTGTGGAATATGAAAAGGACACACTTAAG 1500
QY 1501 aaaaagaataaatacaagtgtagatcccttccctgttcttgcaagtgtagtgaaag 1560
Db 1501 AAAAGATTAATTAAGTATGATGATCCCTTCTCTGTTCTGCAAGTGTCCGAAAAAG 1560
QY 1561 actggaacaaaggaagaaagtgagattgcaacatattctcagaaagattgacaggaagat 1620

Query Match	97.48	Score 2648	DB 9	Length 2780
Best Local Similarity	97.88	Prod. No. 0		
Matches 2719	Conservative	0	Mismatches	0
			Indels	61
			Gaps	1
Qy	1	aattcgcaagactctctcgtgagctgycgccaagacttgaagcggcgagacagagccgg	60	
Db	1	AATTTGGGACACAGCTCTCTGAGGCTTCGGCAAGACTCTGAAGCGCGGACCGAGACCCGG	60	
Qy	61	gctctagaactttagaagacaacggaaatgtaagcgcgggttagagggggaacacaaactctgag	120	
Db	61	GTCTAGACTAGAGACCAACGGAAATGAGCGCGGGGTATAGCGGGAAACCAACTCTGAG	120	
Qy	121	ggccgaagcggagcgcgagaaagaaacggcgcgctgaaggggcggggagcgcaacgaa	180	
Db	121	GGCCAGACGAGCGCGGAGAAAGACCGCGCTGAGGGGGCGGCGCCACCGGAA	180	
Qy	181	ggcaagcgcaggggcgaagcgcgagagcggggagacgcgctggcgctgaagacgggacg	240	
Db	181	GGGAGCGACGAGGGGGCAGCGCGGAGCGGGGACCGCGCTGGCGGTGAGACGGGCGGG	240	
Qy	241	cggcgatctcgcgcgcgcgcgcgctggaacgctttcaaatcttcaaacccgcgcgcca	300	
Db	241	CGCGGATCTCGCGCGCGCGCGCGCTTGAACGCTTTTCAAACTTCAACCGCGCGACCCA	300	
Qy	301	ctcggtttgctttgaccttcctctcgcgcctttgagcggaatcgcggcccggaac	360	
Db	301	CTCGTTTGTGCTTTGGCCCTTCTCTCGGCGCTTGAGCGGATCCGGATCCGGCCCGGAAC	360	
Qy	361	cgagactgaagacggatgactctactctgtgtagaagccgtgagctcgcgagaaagaaag	420	
Db	361	CGAGACTGAAAGCGGATGACTCTACTCTGTGTAGAAGCCGTATGCTGGGGAGAGAGAGG	420	
Qy	421	cggcgctctgtcaaacagcgcgagggaagacgctgtttcgagcgctcccggtggaact	480	
Db	421	CGGCGCTCTGTCAACAGCGCGGGAAGACCGCTTTCGGCGCTGCCCGGTGCCACACT	480	
Qy	481	tcttcggagcccaagatgtagtgcgcggcgagctgccttgaactccggagacgaatgaat	540	
Db	481	TCTTCGGAGCCCAAGATGTAGTGTGTCGGGCGAGCTGCCGTGAACTCCGAGACATGAGAT	540	
Qy	541	ccacagtaaaaggaacttccagggtggaatccaaagtcaaaaatgaaaaaaacagaccat	600	
Db	541	CCACAGTAAAGGACATTTTCCAGGTTGGAATCCAACTCAAAAATATAAAAACAGACACT	600	
Qy	601	tctggaactctcgaaacctgataacagcgccgaagaaatccaaatgtgaagccattgggg	660	
Db	601	TCTGGAATCTCGAAAACTGATAACAGCGCCAGAAAAATCCAAATGTGAAGCCATTTGGGG	660	
Qy	661	aaaagaatttcccttgaacttctcttgcacatatctgaanaacttcaaaaggaat	720	
Db	661	AAAAGATTTTCACTTGACTTACCTTCTGTGCACATATCTGAAAAAATTAAAAAGGACAT	720	
Qy	721	taagagctcgggaagggcaggttgaagaattctcagcaaaagatacagttatcttaattc	780	
Db	721	TAAAGACTCGGAGGGGAGGTGAAAGAAATTTCTGAGCAAAAGATATCAGTATCTTATTTTC	780	
Qy	781	aaataaagaaagaaagtaaatctgcaaaacctggggtcgaaattctctcttacaagatcc	840	
Db	781	AAATAAAGAAAGGATAAATTTGCACAAACCTTGGGTGGAATTTCTCTGTACCAAGTCC	840	
Qy	841	agaatctgcatatctatgtaagaacaaacttcaactatcccgaccgatgtagagttcat	900	
Db	841	AGAACTGCATATCTATGTAAGAACAACCTTCACTTCACTCCAGCCATGATGAAGTTCAAT	900	
Qy	901	taagtaaccaagacaagtggttttaagcagaagaaatattatgttfaaaaagctatcga	960	
Db	901	TAAGTACCAACGACAGTGTGTTTAAGCAGAGAAATATTATTAGTAAAAACTATCA	960	
Qy	961	ggaccatgatattatctctccaatagtatataccaatgcctctgtcagcgggagtaaa	1020	
Db	961	GGACCATGATTTTATTTCCCTCAAAATGATATATTATCAAAATGCTGTGTATGGGAGTAAA	1020	
Qy	1021	aattcttcatattgtagcatttgaacttatacttgaacaaagaaaaaagagttgattt	1080	

Db	1021	ATTCTTCATATGATGACATTTGATAGTACTACATTTGACACAAAAGAAAAAGAGTGTATTT	1080
Qy	1081	accacgaagaaaccaagctctccagctaaagaaatgggggcaaaagatctggtagtggtagca	1140
Db	1081	ACTCAGAAATCAAGTACTCTCAGTAGAGATGGGGCAAAAGATTGGTAGGGGACACA	1140
Qy	1141	aaaaaacaacaagaagactccaaaagcctttgttaaaggtgagaataagaacca---	1197
Db	1141	AAAAACAACAAACGGAAGACTCAAAAAGCCTTTTGTAAAGGTGAAGATATGAGCAACT	1200
Qy	1198	-----aa	1199
Db	1201	TTATAGGCCATTTTCTTCAGCTGACCAATAGCCTTTATATAATTTATTCATTCAGAA	1260
Qy	1200	ggccctcagtcacatttgatctagaacaagccatctgtagtggaaagcaaacctcagttga	1259
Db	1261	GCCTCGACATTCATTGATGTATGAGCACAAGCACTAGTATGCAAAACCAATCAGCTTAA	1320
Qy	1260	actlaagaatccaaacaagatggcgataagtaagtgagaaectcaattcaactccagttgaa	1319
Db	1321	ACTTAGAATCCAAACAGATGGCGATTAAGTATGCTGGAACCTCAATTCATCACTCACTGAA	1380
Qy	1320	agagaagaagaaanaaagatctgtgaaatgttgctcttgcaagaatatagaagactcagaac	1379
Db	1381	AGNAAGAAAGAAAAAGCATTTTGTGAAATGGTTGTCAGAAATATGAAGATCTGAAAC	1440
Qy	1380	tcacacttcaagtgagcaacacagaacactttgcacaagatataccagatatacagttgttga	1439
Db	1441	TCACCTTCTTAAGTGAACACACAGAAACTTTCACAGAGTAAACAGATCAAGTGGTTGA	1500
Qy	1440	tgaatctgatactcaagttagttcttctgactcttggaataatagaagagacacacataaaaa	1499
Db	1501	TGATATTTGTTCAAGTATAGTTTGTGACTTTGTGAAATGAAAGACACACTTAATAA	1560
Qy	1500	gaagaagaataaaatcaatgttggatcccttctctctcttctctgcaagtgatccttgaanaa	1559
Db	1561	GAAGAATATTAATACAGTGGTGAATCCCTTCTCCTGTTCTGCAAGTGTCTCTAANA	1620
Qy	1560	gacgagacaaagaagaagaatggagatgtgcacacttcttcgaagaagaattggccagaaga	1619
Db	1621	GATCGAACAAAAGAAAAAGTGAATTCGCAATATTTCTCAGAAATATTCGCAGAGAGA	1680
Qy	1620	tgaatacaagtgagaagggagaaattcttcgtgttaaagagccagagaacctgaaaaaa	1679
Db	1681	TGATACACACGTGAAGACAGCAATTTTCGTATAAGAGACCCAGGAACCTGAANAANA	1740
Qy	1680	gtccctcttattatagaagccatcccccaccccttcaaatgaatcttggagggcttaaga	1739
Db	1741	GCTCTGTATTATTTCAGACCCCATCCCCACCTTCMAATGAATTGAGAGGCTTAATGA	1800
Qy	1740	gaaaaagagtaataaattgtccatggttaagaagtgagaatgagaaatgaagaagaattt	1799
Db	1801	GAATAATAGTAATAATGTGCAATGTATATATACGTACAGTACATTAAGACAGAAATT	1860
Qy	1800	tacacagctacactctacataaaacaacaggaatgcattcttgacacttcgaaacacac	1859
Db	1861	TACACACCTACCTTACATATAAAACAAACAGATATCATTTCTTGACTTTCCACACAC	1920
Qy	1860	attaagtgaaatagacttangaagaactaaggtagatacactataaagtatacaatcacgc	1919
Db	1921	ATTTAAGTGAANAATGACTTGAAGAACTTAAGCGTAGTCTATTAATATGTAACTACAGGC	1980
Qy	1920	atctgtacatgttctcgatcttaagtlacagataaagtgatcttcaaccacaagaagtc	1979
Db	1981	ATCTGTACATGTTTCTGATTTAGTACAGATTAATAGTGCATCTCAACCAAAACAGAAATC	2040
Qy	1980	agataacgtgctttcttcagcaaaagatctccaagagaagaagagccctcaactcaatctac	2039
Db	2041	AATATCTGTGCTTTTCCAGCAAAAGATCTCAAGAAAGAGCACTTCATTCATTAATTTCAC	2100
Qy	2040	tcattgattctggctgtatacaataaacaagtttcaacaagagccatactgttcaggcaaa	2099

Db	2101	TCATGATTCCTGCTGTAATCAATAAAGACTTTCACAGAGACCTAATCTTTCAGGCAAA	2160	CDS	217..2241	/note="HsCdc7 protein kinase regulatory subunit"
Qy	2100	ggctcattccactactccctcctgaggaaacccaatgaatgagctcgaataatgagatg	2159		/codon_start=1	/product="Dbf4 protein"
Db	2161	GGCTCCATTCCATCTCCCTCCGAGAGACCCAAATGAATGTGACTTCAAGAAATATGATGATG	2220		/protein_id="AAD41911.1"	/db_xref="GI:5306125"
Qy	2160	ttctactctctgttaaatatcatcagaagaatgaaataatattagagcgaatacagaaga	2219		/translation="MNSGAMRHSKGFQGGIOWNEKNRPSLKLKTDNRPESKCK	/translaltion="MNSGAMRHSKGFQGGIOWNEKNRPSLKLKTDNRPESKCK
Db	2221	TTTACTCTCTGTAAATAATCATCGAAGAAATGAAATTAATTAAGACGAAATATGAAAAA	2280		PLMKRVDLPSVTSISELQKDIKDLGGAGFEELSKDIYLSIKRKAADKDDFTPSNIIIS	PLMKRVDLPSVTSISELQKDIKDLGGAGFEELSKDIYLSIKRKAADKDDFTPSNIIIS
Qy	2220	aaactcggaaacaaatgctgtaatttgataaagaactgaaattatatacaagaagaana	2279		NALSMGVAILHDDIRITYIEOKRKLVLKSSISVRDGRKRVSGAKRTIRGRKPK	NALSMGVAILHDDIRITYIEOKRKLVLKSSISVRDGRKRVSGAKRTIRGRKPK
Db	2281	AAATCTGGAACCAATAGCTGATTTGATTAAGAAGAACTGAATTTATATACAGAAAGAAAA	2340		FVKVMDMSQLRPPYLQJLTMPEFINVSIQKPCSPDVDPKPSMOKQOTGVKRLIQDGD	FVKVMDMSQLRPPYLQJLTMPEFINVSIQKPCSPDVDPKPSMOKQOTGVKRLIQDGD
Qy	2280	cagaattctgtatgtaaccggtacagctcttactagaacttgcttcagactagtgagaaga	2339		KYGGTSIOLQKEKKKKGCECCLOQYEDLETHLSEQRNEAQNQOYVDIVYSLK	KYGGTSIOLQKEKKKKGCECCLOQYEDLETHLSEQRNEAQNQOYVDIVYSLK
Db	2341	CAGAAATTTGTAATTCACCGGTACAGCTTTTACTAGACTTTTCAGACTTTCAGAGAGAA	2400		VEDPEYFKDTPPKKRIRKYSVGSLSPEVASYLKTEQKEKVELHISQKDCQEDDTTV	VEDPEYFKDTPPKKRIRKYSVGSLSPEVASYLKTEQKEKVELHISQKDCQEDDTTV
Qy	2340	atcagaattcttggtcttcacaagctacaagaagaagtggtatataatgatttctttaga	2399		KEONFLKETQETEKKLIFSEPIPHSPNENLGRNEMKSNMSTAEEDIRQWFO	KEONFLKETQETEKKLIFSEPIPHSPNENLGRNEMKSNMSTAEEDIRQWFO
Db	2401	ATCAGAAATTTTGGGCTTTCACAGACTACACAGAAAGAGTGGTATATGCAATGTTTTAGA	2460		LPLMKRQECILDISEHRLSEHDELELVYDRIKCTQASVHVSDETSNSSQPKOKS	LPLMKRQECILDISEHRLSEHDELELVYDRIKCTQASVHVSDETSNSSQPKOKS
Qy	2400	tatttgggaagaagaataatcagataatcgttaacagcgttttctcgtccctcaac	2459		DVLEPANDLKEKDLHSITFTDSGLTITINSQSEHLTVQAKAFEPHPEPEPCDKNM	DVLEPANDLKEKDLHSITFTDSGLTITINSQSEHLTVQAKAFEPHPEPEPCDKNM
Db	2461	TATTTGGGAGAGGAAATTCAGATATCTGTTAAACAGCGTTTTCCTCCCTTCAAC	2520		DSLPSGKILHRVKYIILGRNREKNLEPNAEFQKRTFELIODEBNRILCSSVQSIILDFOT	DSLPSGKILHRVKYIILGRNREKNLEPNAEFQKRTFELIODEBNRILCSSVQSIILDFOT
Qy	2460	tctctcaattactgcttctttagaatttaaaaaatgcatccttccagaagtgataagat	2519		SEEKSEFLGFTSYTERKSGICWVLDIWEENSNDLLTAFPSSTSTFGF"	SEEKSEFLGFTSYTERKSGICWVLDIWEENSNDLLTAFPSSTSTFGF"
Db	2521	TTCTCATTTTACTGGCTTTTACAAATTTTAAAAATGATGATCTTTTTCAGAGATTAAGAT	2580			
Qy	2520	catattcttgaattttttataatctgtatgaaatctcagaattcttccagctt	2579			
Db	2581	CATATTTCTTGAATTTTATTAATATGATVAGAAATCTTACGATTTTTCACCATCTTT	2640			
Qy	2580	gtttacagagcccaatgtaaatatataataataatatttgcaatttctcagaatttga	2639			
Db	2641	GTTTACAGACCAATGTAATATTTAAAAATTAATTTTGCAATTTTTCACAGAAATGGA	2700			
Qy	2640	tacctgtttaaagaataatacagaataaactgtgactgtctgttcttacttaaaaa	2699			
Db	2701	TACCTGTAAAGAAAAAATACAGAAATTAACCTGTGACTGCTGTTTACATTAAAAA	2760			
Qy	2700	aaaaaaaaaaaaaactcgag	2719			
Db	2761	AAAAAAAAAAAAAACTCGAG	2780			
RESULT	3					
AF160249	2474 bp	mRNA	PRI		01-JUL-1999	
DEFINITION	Homo sapiens Dbf4 protein mRNA, complete cds.					
ACCESSION	AF160249					
VERSION	AF160249.1	GI:5306124				
KEYWORDS						
SOURCE	human.					
ORGANISM	Homo sapiens					
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.					
REFERENCE	1 (bases 1 to 2474)					
AUTHORS	Jiang, W. and Hunter, T.					
TITLE	Mammalian Cdc7/Dbf4 Protein Kinase Complex Is Essential for					
	Initiation of DNA Replication					
JOURNAL	Unpublished					
REFERENCE	2 (bases 1 to 2474)					
AUTHORS	Jiang, W. and Hunter, T.					
TITLE	Direct Submission					
JOURNAL	Submitted (17-JUN-1999) Molecular Biology and Virology Laboratory,					
	The Salk Institute, 10010 North Torrey Pines Road, La Jolla, CA					
	92037, USA					
FEATURES	Location/Qualifiers					
source	1..2474					
	/organism="Homo sapiens"					
	/db_xref="taxon:9606"					


```
/db_xref="taxon:9606"
/chromosome="7"
/map="7q21"
1..2460
/gene="DBF4"
200..2224
/gene="DBF4"
/codon_start=1
/product="DBF4-like protein"
/protein_id="A04537.1"
/db_xref="GI:5566281"
/translation="MNSGAMRIHSKGFHOGIOVKNENRPSLKLTKDNPEKSKC
PLMKVFLDPSVTSISKLOKIDKDGREELSLKISLISKKAKAKAQTIGRI
SPVSPSATAETSPHSDSSFPVCSLRGLLYEKAKEHDFIPSNILS
NALSQVLIHIDIRYIEOKKELYLKSSYVRDGRVSGAOKRTGRILKP
FVKEVDSQILRYPIQLTNMPLINYSLQKSPVDKPPSMQKQIVKLIQTGD
KYGGTSIOLQLEKKKKGYCECCLOKYEDELTHLSEDRHFAVSNQVVDIYSKL
VFDEVEKEDTPPKRKRIKYSVGLSPVSAVLKKEOKEVELDHSIISOKDOEDTV
KEONFLKEQETEDSKLFTSEPIPHPSNELRKNENKSNLSMSTAEEDIRFOTQ
LPLHKNQECIDISEHTLSNDLELRVDYKKNQASVAVSDPSDINSQSPKOKS
DTVLPPAKDLEKDLHSLFTHDSGLITINSQEHLYQAKAPFHPPEEPNCDKKN
DSLPSGIHRKVKIILGRNRKENLEPNAPFPRKREFLTQENRICKSPVQSLDLEOT
SEKSEPLGFSTYKSGICNVNDIWEENSDNLTAFSPSPSTSTGF"
```

BASE COUNT 891 a 445 c 475 g 649 t

ORIGIN

Query Match	85.5%	Score 2324	DB 9	Length 2460
Best Local Similarity	97.5%	Pred. No. 0		
Matches 2395	Conservative 0	Mismatches 0	Indels 61	Gaps 1

```
QY 319 ctctctctccgcgcttggagcgagatccgccccggaaaccgacctgcagagcgagt 378
DB 1 CTTCTCTCTCCGCGCTTGGAGCCGGAATCCGCGCCCGGAAACCGACCTGCAGACGCGGT 60
QY 379 acctactcgtlaaagccgttagctggcgaagagagagagcgcgccgtctgcaag 438
DB 61 AACTTACTGCTAAGGCCCTAGCTGGGGAAGAGAGAGAGCGCGCTCTGTCAACAG 120
QY 439 gccgggggaagccgtgtcttcggcgctgcccggtgcagacattctccggaccagcaty 498
DB 121 GCCGGGGGAAGCCGTGTTCGCGGCTGCCGCGGTGCAGACATTTCTCCGAGCCACAGCATG 180
QY 499 taggtgcgggggactgcatgaactccggagccatgtgagatcccaagtaaaaggacatt 558
DB 181 TAGGTGCCGGGGAAGTGCATGAACTCCGAGACCATGAGATCCACAGTAAGGACATTT 240
QY 559 ccaagggtggaatccaagtcanaaatgnaaaaaacagacacatctcgaatctctgaaac 618
DB 241 CCAGGCTGGAATCCAAAGTCMAAAATGMAAAACAGACCATCTCTGAAATCTCTGMAAAC 300
QY 619 tgaatacagcgccagaaaaatccaatgtaagccactttggggnaaaagtatttacctga 678
DB 301 TGAATAACAGGCGCGAATAATCCAAATGTAAAGCCACTTTGGGGAATAAGTATTTACTCTGA 360
QY 679 cttaacctctgcacatatctgtaaaacttcaaaaggacataaaggactctggggggg 738
DB 361 CTTAACCTCTGCACATATCTGAAAACCTTAAAAGGACATTAAAGATCTGGAGGCGG 420
QY 739 agtgaagaatttccagcaaaagatacagttatcttatttcaaaataaagaagaagctaa 798
DB 421 AGTGAAGAATTTCAGCAAAAGATATCAGTTATCTTATTCAAAATAAAGAAAGACTAA 480
QY 799 atttgacaacacttggtgcgaattctctcgttaccaagtcagaactgcatactgc 858
DB 481 ATTTGACAAACCTTGCGTGAATTTCTCTGTACCAAGTCCAGAAATCTGCATATTTACTGC 540
QY 859 agaaacccttcaactctccagccatgatggaagttaattaaagtaaccgcaaacagt 918
DB 541 AGAAACCACTTCACTCATCCAGCCATGATGAAAGTATTATTAACTCACCGACAGCACT 600
QY 919 ggttttaagcagaagaaaaatlaagttgaaaaagctatcaaggacagatatttattcc 978
```

```
DB 601 GGTGTTTAAGCAGAGAAAATATATAGTGAAGAAAGCTATCAAGGACCATGATTTATGCC 660
QY 979 ttcaaatagatataatacaatgacctgtgcattggggagtaaaaaattctcatattgata 1038
DB 661 TTCAATAGTATATATCAAAATGCCCTTGATCGAGGAGTAAAAATCTTCATATTTGATGA 720
QY 1039 catltagatcatcatattgaaacaaaagaaaagatgttatttcaataaagaatcaagtttc 1098
DB 721 CATTTGATATCTACATTTGAACAAAAGAAAAGAGTGTGATTTATCTCAAGAAATCAAGTAC 780
QY 1099 ttcaagtaagatggygggcaaaaagttgtagtgggtgcacaaaacagaacagaagaag 1158
DB 781 TTCAGTAAGATGGGGGCAAAAAGATTGGTAGTGTGCACAAAAACAAAGACAGAAG 840
QY 1159 actcaaaaagcctttgttaaagggtggaatatgagcca----- 1197
DB 841 ACTCAAAAAGCCTTTGTAAAGGTGAAGATATGACCAACTTAAAGCCATTTATCT 900
QY 1198 -----aagccctgagttccattga 1217
DB 901 TCAGCTGACCAATATGCCCTTTTAAATTAATTTCTATTCAGAAAGCCTTGACATCTATGGA 960
QY 1218 ttgaagcaagccatctagatgcaaaagcaaaactcaagttaaactaagaatccaacaga 1277
DB 961 TGTAGCAAGCATCTATGATATGCAAAAAGCAACTCAGTTAAACTAAGATCCAAACAGCA 1020
QY 1278 tggcgataaglatggttgaacctcaattcaactcaagttgaaagagaagaagaagaag 1337
DB 1021 TGCAGTAAGTATGGTGGGAACCTCAATTCACCTCAGTTGAAAGAAAGAAAGAAAAAG 1080
QY 1338 atatttgaaatgttgcttgcagaataataaagatcagaagaactcactcttaagttaga 1397
DB 1081 ATATTGTGAATGTGCTTCAGAAATATGAAGATCTAGAAACTCACCTTCTAATGATGGA 1140
QY 1398 acacagaacttgcacaagaagtaaccagatcaagttgttgaatcgtatcattgaagt 1457
DB 1141 ACACAGAAACTTTGACACAGATATACAGATATCAAGTTGTTGATGATATGATCTAAGTT 1200
QY 1458 agctttgaacttgggaatagaaagacacacactaaagaagaagaataaatacag 1517
DB 1201 AGTTTGTGACTTGTTCGAATATGAAAAGACACACTTAAAGAAAAGAAATTAATAACAG 1260
QY 1518 ttttgatcccttctccgttctgcgaagtgtccgtgaaaagaactgaaacaaaagaaaa 1577
DB 1261 TGTGTGATCCCTTCTCTCGTTCGATTCGAAAGTGTCTGAAAAGACGTGAACAAAAGGAAAA 1320
QY 1578 agtgaattgcaacatatattctcaagaagattgcccaggaagatgatatacaacagttgaaga 1637
DB 1321 AGTGAATTTGCAACATATTTCTCAGAAAAGTTGCCAGGAAGATGATCAAACTGAAGGA 1380
QY 1638 gcaagaattccgtatataaagagacccaggaacatgaaaaaagctcgtttatttcaga 1697
DB 1381 GCAGATTTCTCTGTATTAAGAGACCCAGGAAACTGMAAAACCTCTCTGTTTATTTTCAGA 1440
QY 1698 gcccatcccccacccttcaaatgaaattgagaagggttaattgaaaataatgaaataaag 1757
DB 1441 GCCCATCCCCACCCTTCAAAATGAATTTGAGAGGCGCTTAATGAAAAATGATGATTAAG 1500
QY 1758 ttccatgtttaaagtagcgtgaagatgacataaagaacagaaatttatacaagctatacttaca 1817
DB 1501 TTCCATGTTAAAGTAGACGTGAAGATGACATAAGAGAAATTTTACACAGCTACTCTACA 1560
QY 1818 taaaaaacaacaggaatgcattcttgacattccgaaacacacataaagtgaagaatgact 1877
DB 1561 TAAAAACAAACAGGAATGCAATCTTGACATTTCCGAACACACTTAAGTGAAGAAATGACT 1620
QY 1878 agaaagactaaagtagatcacatataatgtaacataaaggaacatcgtatcatattctga 1937
DB 1621 AGAAGACTTAAGGATGATCATTATTAATGTATACATACAGCGCTCTGATATTTCTGA 1680
QY 1938 ttccagtaagataatagttgattctcaaccacaaagaagtcagatactgtgtttcc 1997
DB 1681 TTTCACTACAGATATATAGTGAATCTCAACCAAAAAGAAAGTGAATAGTGTGTTTTC 1740
```

QY 1998 agcaagaagatctcaaggaagaaagacccttcacatattacacatgattcgtcgtgat 2057
 |||||
 DB 1741 accaagaagatctcaaggaagaaagacccttcacatattacacatgattcgtcgtgat 1800
 QY 2058 aacaaataaagatctcaaggaagaaagacccttcacatattacacatgattcgtcgtgat 2117
 |||||
 DB 1801 aacaaataaagatctcaaggaagaaagacccttcacatattacacatgattcgtcgtgat 1860
 QY 2118 tccgaggaagacccttcacatattacacatgattcgtcgtgat 2177
 |||||
 DB 1861 tccgaggaagacccttcacatattacacatgattcgtcgtgat 1920
 QY 2178 acacgaaagatctcaaggaagaaagacccttcacatattacacatgattcgtcgtgat 2237
 |||||
 DB 1921 acacgaaagatctcaaggaagaaagacccttcacatattacacatgattcgtcgtgat 1980
 QY 2238 tgaattgataaagaaatgaattatctacacagaaagaaagaaatgattgattcacc 2297
 |||||
 DB 1981 tgaattgataaagaaatgaattatctacacagaaagaaagaaatgattgattcacc 2040
 QY 2298 ggtacagatcttactagatgattcgtcgtgatgaaagaaatcagaatttctggtt 2357
 |||||
 DB 2041 ggtacagatcttactagatgattcgtcgtgatgaaagaaatcagaatttctggtt 2100
 QY 2358 caaaagctacacagaaagaaagatgattcgtcgtgatgaaagaaatcagaatttctggtt 2417
 |||||
 DB 2101 caaaagctacacagaaagaaagatgattcgtcgtgatgaaagaaatcagaatttctggtt 2160
 QY 2418 ttaagtaattcgttaagaaagatgattcgtcgtgatgaaagaaatcagaatttctggtt 2477
 |||||
 DB 2161 ttaagtaattcgttaagaaagatgattcgtcgtgatgaaagaaatcagaatttctggtt 2220
 QY 2478 ttaagtaattcgttaagaaagatgattcgtcgtgatgaaagaaatcagaatttctggtt 2537
 |||||
 DB 2221 ttaagtaattcgttaagaaagatgattcgtcgtgatgaaagaaatcagaatttctggtt 2280
 QY 2538 ataaatcgtatgaaagaaatcgtatgaaagaaatcgtatgaaagaaatcgtatgaaagaaat 2597
 |||||
 DB 2281 ataaatcgtatgaaagaaatcgtatgaaagaaatcgtatgaaagaaatcgtatgaaagaaat 2340
 QY 2598 aatattcgtatgaaagaaatcgtatgaaagaaatcgtatgaaagaaatcgtatgaaagaaat 2657
 |||||
 DB 2341 aatattcgtatgaaagaaatcgtatgaaagaaatcgtatgaaagaaatcgtatgaaagaaat 2400
 QY 2658 taaagtaattcgttaagaaagatgattcgtcgtgatgaaagaaatcagaatttctggtt 2713
 |||||
 DB 2401 taaagtaattcgttaagaaagatgattcgtcgtgatgaaagaaatcagaatttctggtt 2456
 RESULT 5
 AC018977/c 236236 bp DNA HTG 26-JUL-2001
 LOCUS Homo sapiens chromosome 10 clone RP11-386C23, WORKING DRAFT
 DEFINITION
 AC018977.6 GI:15021987
 ACCESSION
 AC018977.6 GI:15021987
 VERSION
 AC018977.6 GI:15021987
 KEYWORDS
 HUG, HTGS_PHASE1, HTGS_DRAFT, HTGS_FULFILL, HTGS_ACTIVEFIN.
 SOURCE
 Human.
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 REFERENCE
 1 (bases 1 to 236236)
 Smith, D.R.
 TITLE
 Genome Therapeutics Corporation Sequencing Center: Human Genome
 Sequence Data
 JOURNAL
 2 (bases 1 to 236236)
 Smith, D.R.
 AUTHORS
 Direct Submission
 TITLE
 Submitted (25-DEC-1999) Genome Therapeutics Corporation, 100 Beaver
 JOURNAL
 Street, Waltham, MA 02453, USA
 COMMENT
 On Jul 26, 2001 this sequence version replaced gi:14091812.

 Genome Center
 Center: Genome Therapeutics Corporation
 Center code: gtc
 Web site: http://www.genomecorp.com/
 Contact: gtc-seqcenter@genomecorp.com

 Project Information
 Center project name: hg92

 Summary Statistics
 Sequencing vector: N/A
 Chemistry: Dye-terminator Big Dye; 1008 of reads
 Assembly program: Phrap; Version 990315
 Consensus quality: 228735 bases at least Q40
 Consensus quality: 230209 bases at least Q30
 Consensus quality: 231169 bases at least Q20
 Insert size: 235235; sum-of-coverage
 Quality coverage: 6.3x in Q20 bases; sum-of-coverage

 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 12 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.
 *
 1
 1147: contig of 1146 bp in length
 1247: gap of unknown length
 1447: contig of 1220 bp in length
 1467: gap of unknown length
 2467: gap of 1459 bp in length
 2567: contig of 1459 bp in length
 4026: gap of unknown length
 4125: contig of 1040 bp in length
 4126: gap of unknown length
 5166: gap of 1044 bp in length
 5266: contig of 1044 bp in length
 6309: gap of unknown length
 6410: contig of 1143 bp in length
 7552: gap of unknown length
 7553: gap of 1064 bp in length
 7653: contig of 1064 bp in length
 8717: gap of unknown length
 8817: contig of 1491 bp in length
 10308: gap of unknown length
 10408: gap of 1506 bp in length
 11914: contig of 1506 bp in length
 12013: gap of unknown length
 12014: contig of 1350 bp in length
 13364: gap of unknown length
 13364: contig of 9324 bp in length
 106707: gap of unknown length
 108708: contig of 129429 bp in length.
 108808: Location/Qualifiers
 FEATURES
 source
 1.236236
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="10"
 /clone="RP11-386C23"
 /clone_1fb="RP11-11"
 1.1146
 /note="assembly_name:Contig7"
 1247.2466
 /note="assembly_name:Contig11"
 2567.4025
 /note="assembly_name:Contig12"
 4126.5165
 /note="assembly_name:Contig13"
 5266.6309
 /note="assembly_name:Contig16"
 6410.7552
 /note="assembly_name:Contig17"
 7653.8716
 /note="assembly_name:Contig18"
 8817.10307
 /note="assembly_name:Contig20"
 10408.11913
 /note="assembly_name:Contig21"

[illegible][illegible]


```

Oy 1110 atggggcacaagaagtgtgtgtgtgtgcacaaaacacagaaggaactcaaaaagc 1169
    || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 787 ATTCGGGAAAAAGAGCTGGCAATACATAAAGGCGAGAGAGACTCAAAAGG 846
Oy 1170 ctttctaaaggtggaagataatga----- 1193
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 847 CTTTAAAAAGGTGAAGATGGAACCGGCTCTACAGGCCATTCCTTACGTCAGCA 906
Oy 1194 -----gccaaagccctgcagtcacattgagtgagacaagc 1228
    || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 907 GTGGCCCTTCTATAACATGCTACGACAAAGCCCTTGACGCCATTTGACATAGAGAAC 966
Oy 1229 catctatgatgcaaaagaacactcaaggttaaaactaagaatccaaacagatgcaagt 1288
    || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 967 CATCCAGGCTCAAAAGCAACCTCAGCCCTAAACCCAGCTAAATAGGATGTTGAAT 1026
Oy 1289 atggtggaacctcaattcaactccagttcgaagaagaagaagaagaagaagataatg 1348
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1027 GTGGTGAACCCCAAGTTCATCTCCATTTGAAGAAAGAAAGAAAGAAAGATTTGAAAT 1086
Oy 1349 gtgcctgcagaataatgaagatctagaactcaactctaagtgagacaagacaagaa 1408
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1087 GTTCTTACAGAAATACGAAGATCTTGAACCTTTCAGTTCAGTGAAGAACCAAAACT 1146
Oy 1409 ttgcacagagtaaacacagatcaagttgtgtatgtatgtatcttaagtttaagtt 1468
    || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1147 TCGCAACAGATACCAAGATTCAGATGTTGATGATGATGATGATGATGATGATGATG 1206
Oy 1469 ttgaggaatctgaaagaacacacactaaagaagaagaagaagaagaagaagaagtt 1528
    || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1207 TTGTGGAATATGAGAGAGACACACCTTAAAGAAAGAAAGAAATATACGCGTTGCTCT 1266
Oy 1529 tttctcctgttctgcagaagtgtcctcgaagaagactgcagaagaagaagaagtg 1588
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1267 TTTCTTCTGTTTTCGAATGTCTCTGAAGAAACACTGAAACCAAGAGAAAGTACAGTTGG 1326
Oy 1589 aacatattctcagaagaatgtccagaagaagatgatacaaacagtgagaagacaga----- 1642
    || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1327 AGCCCATTTTCCAGAGGACATGTGTAAGATTAATGACCACTGCCCTGAGGATTTTTC 1386
Oy 1643 -----attcctgtataaagagacacaggaacacgtaaaaaagctccgttatatt 1693
    || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1387 AGTGGAGATATTCAGTGTGAAGTATCCAGAAACCTGGAACAAAGGCTTGTGTTGCCCT 1446
Oy 1694 cagagcccatcccccaccccttcaaatgtaattgagagggcttaattgagaataatg 1753
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1447 CAGAACCCATGTCTACTCTTCAACGGGATTTGAAGGACGATGATGAGAAA-----G 1497
Oy 1754 aatgttccagttaagtaacagctgaagaatgacataagaacagaattttaacacagctactc 1813
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1498 CAGCCAGCATGTTAAATGCCAGTGAACCTGACATPAAACAGAAAGTTTACACAGTACCTC 1557
Oy 1814 tacataaaaaacacagaagaatgcatcttgcataatccggaacacacataagtgaaagt 1873
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1558 CATGTAATAAATCAAGAGAGAAATTTCTGATCTTCAACATTAATTAATTAATA 1617
Oy 1874 acttga-----agaactaagggtagatgacttaactaaatgtaacatacagcattgacatg 1930
    || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1618 GAATGATTTAGACAGAGGTTGTTGACTCTGTAGGTGTGCCACGGCTCATGTGTACAAAG 1677
Oy 1931 ttctcgtatctcagtaacataatagtgatctcaacacaaacagaagtaacagatctg 1990
    || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1678 TTTCTCATCTCAGTCCAGAAATAGTCTACCCCAACCAAAATTAACAGAGTACTACAC 1737
Oy 1991 ttttcccgcaagaagatctcagaagaagaagacccatcattcaattactatctgattctg 2050
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1738 ATTCTTTCAGCAAAAGATTTCAGAAAGAGACCTTCATTTTGTGTTGCTCATGATTCG 1797
Oy 2051 gctcgtataacacagcttcaacagaagccactacgttctcagaagaagctctcc 2110
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1798 ATCTGTGATGATTAATATACATCAAAAAGAACAGCTAAACAGTGAAGCCGACGCCATCT 1857

```

```

Oy 2111 atactcctctcgtgaggaaacccaatgtaattgactcaagaataatgataagttactctg 2170
    || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1858 GTGCTCTC---AGCAGCCCAATGAGTGTGACACGAGAGATACATTAACCTTGCTGTG 1914
Oy 2171 gtaaatatcatcgaaagtgaaataaataatagacagaagaagaagaagaagatctcgagc 2230
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1915 GTAATATCAGCGGAAAGAGAGAGATGTTATTTAGGGCAGAG---AAAAAATATGTGAGCC 1971
Oy 2231 caaatgctgaatttgatc---aaaagaactgaatttattacacagaagaagaagaagatt 2287
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1972 CAAGTCTGAATTTGATTAATAAAGAAAGACGATGATTTCTCTATGTGTGAAGCAACATTT 2031
Oy 2288 gtagttcccggtacagctcttcaagctgcttctcagactggtgaagaagaataagaat 2347
    || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2032 GTGGGTACCACTACAGTCTCTACGTGACCTTTTCAACAAAGCCGGGGAATATGATTT 2091
Oy 2348 ttlttggttccacagctacacagaagaagtgatataatgcaatgttttgaatcttg 2407
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2092 TTTTGGGTTTACACAGTACACTGAAACAGTGGGCTATGATGTTTATGATTTGGG 2151
Oy 2408 aagaggaatctcagataatctgttaacagcglttctcgtccctcctaactcaat 2467
    || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2152 AGGATGAATAATTCAGTATGCTGCTCAACATTTTCTCTCCCTCAGCTCTACAT 2211
Oy 2468 ttactgctttagaatttaaaaaatgcatactttcagaagtgatgaatcatattct 2527
    || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2212 TCATTGGAATTTATGATTTAAATGTTGTTTTCAGAAAGTACAGTTTATTTCTGGAATTTT 2271
Oy 2528 tga-----aatttataataatgatatgaaatcttgaattcttaccag 2575
    || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2272 TATTAATACATATATATATTTCTTAGAGAGTCTGCTTTTGGTTTGGTTGTTTCCGAG 2331
Oy 2576 ctgttaccagaccacaaatgtaaatatataaataaataatgtaattcttaccagat 2635
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2332 CTTTGTGCTAGAACCATATGTAATTAATAAGAA--TGATGTTTATTAATTTTAAAGAAAT 2389
Oy 2636 tgaataacctgtttaaagaagaatctacagaataaactgtgacggtctgttttaccataa 2695
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2390 TGAGTAGCTGTT--ACAAATTAATCAATTAATAATCTGTGATTTATTTATTAATCAATCAAAAAA 2447
Oy 2696 aaaaaaaaaa 2706
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2448 AAAAAAAAAA 2458

```

RESULT 7
AC005164/c
LOCUS AC005164 123331 bp DNA PRI 03-FEB-2000
DEFINITION Homo sapiens BAC clone CTB-135C18 from 7q21, complete sequence.
ACCESSION AC005164
VERSION AC005164.1 GI:3242749
KEYWORDS HTG.
SOURCE human.
ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Kellen,J. and Burkhardt,J.
TITLE The sequence of Homo sapiens BAC clone CTB-135C18
JOURNAL Unpublished
REFERENCE
AUTHORS 2 (bases 1 to 123331)
 Waterston,R.
TITLE Direct Submission
JOURNAL Submitted (20-JUN-1998) Department of Genetics, Washington
 University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
 3 (bases 1 to 123331)
REFERENCE
AUTHORS Wacerston,R.
TITLE Direct Submission
JOURNAL Submitted (03-FEB-2000) Department of Genetics, Washington
 University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
COMMENT ----- Genome Center
 Center: Washington University Genome Sequencing Center
 Center code: WUGSC

Web site: <http://genome.wustl.edu/gsc>
 Contact: sapiens@wustl.wustl.edu
 ----- Summary Statistics -----
 Center project name: H_RG135C18

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded or sequenced with an alternate chemistry; an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:

The sequence of this clone was established as part of a mapping and sequencing collaboration between the NHGRI Chromosome 7 Mapping Project (Eric D. Green, Director), John D. McPherson in the Department of Genetics (Washington University), and the Washington University Genome Sequencing Center. For additional information about the map position of this sequence, see <http://www.nhgri.nih.gov/DIR/CTB/CHR7>, send <mailto:egreen@nhgri.nih.gov>, or see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:

Clone CTB-135C18 is from the first release of the human BAC library CTR-978SK-B. The library contains cloned DNA from the male fibroblast cell line 978SK. See: Shizuya et al., Proc. Natl. Acad. Sci. USA 89:8794-7 (1992); U-J. Kim et al., Genomics 34:213-8 (1996). This clone is available from Research Genetics, Inc. (<http://www.resgen.com>).

VECTOR: pBeltoBAC11
 Selection: chloramphenicol

NEIGHBORING SEQUENCE INFORMATION:

The clone sequences to the left is RP5-1189K17; the clone sequenced to the right is CTB-60N22, 200 bp overlap. Actual start of this clone is at base position 1 of CTB-135C18; actual end is at 123138 of CTB-135C18.

Location/Qualifiers

1.123331

/organism="Homo sapiens"

/db_xref="taxon:9606"

/chromosome="7"

/map="7q21"

/clone_id="CTB-135C18"

/clone_id="CTB-978SK-B"

36.78

/rpt_family="MER2_type"

235.322

/rpt_family="MER2_type"

1466.1636

/rpt_family="Alu"

1995.2122

/rpt_family="MIR"

2782.3080

/rpt_family="Alu"

3083.3134

/rpt_family="MALR"

3136.3664

/rpt_family="Retroviral"

3786.3835

/rpt_family="MALR"

3891.4128

/rpt_family="MALR"

4476.4714

/rpt_family="MIR"

4882.5225

/rpt_family="MALR"

repeat_region 5762.5939
 /rpt_family="MER1_type"
 repeat_region 5950.6035
 /rpt_family="MER1_type"
 repeat_region 6360.6642
 /rpt_family="Alu"
 repeat_region 7010.7308
 /rpt_family="Alu"
 repeat_region 9806.10112
 /rpt_family="Alu"
 repeat_region 10303.10498
 /rpt_family="L1"
 repeat_region 10507.10807
 /rpt_family="Alu"
 repeat_region 10861.10978
 /rpt_family="L1"
 repeat_region 11533.11836
 /rpt_family="Alu"
 repeat_region 11959.12079
 /rpt_family="Alu"
 repeat_region 13992.14290
 /rpt_family="Alu"
 repeat_region 1571.16139
 /rpt_family="L1"
 repeat_region 16142.16438
 /rpt_family="Alu"
 repeat_region 16500.16779
 /rpt_family="Alu"
 repeat_region 16933.17077
 /rpt_family="L1"
 repeat_region 17850.18392
 /rpt_family="L1"
 repeat_region 18480.18769
 /rpt_family="Alu"
 repeat_region 18908.19366
 /rpt_family="L2"
 repeat_region 21782.21896
 /rpt_family="L2"
 repeat_region 22048.22156
 /rpt_family="MER1_type"
 repeat_region 24873.25250
 /rpt_family="MALR"
 repeat_region 26032.26309
 /rpt_family="L2"
 repeat_region 26342.26817
 /rpt_family="MALR"
 repeat_region 27147.27355
 /rpt_family="Alu"
 repeat_region 27381.29443
 /rpt_family="L1"
 repeat_region 29444.29561
 /rpt_family="Alu"
 repeat_region 31170.31909
 /rpt_family="MIR"
 repeat_region 32332.32720
 /rpt_family="MALR"
 repeat_region 32731.32953
 /rpt_family="Alu"
 repeat_region 33094.33227
 /rpt_family="MER1_type"
 repeat_region 33643.33941
 /rpt_family="Alu"
 repeat_region 34000.34088
 /rpt_family="Alu"
 repeat_region 34233.34602
 /rpt_family="MALR"
 repeat_region 38520.38721
 /rpt_family="MIR"
 repeat_region 38730.38849
 /rpt_family="Alu"
 repeat_region 40214.40605
 /rpt_family="L2"
 repeat_region 41565.41638

```

repeat_region /rpl_family="L1"
42039..42347
repeat_region /rpl_family="MIR"
42470..42784
repeat_region /rpl_family="L1"
42787..43087
repeat_region /rpl_family="Alu"
43588..43892
repeat_region /rpl_family="Alu"
44036..44116
repeat_region /rpl_family="L1"
44562..45028
repeat_region /rpl_family="L1"
45191..51321
repeat_region /rpl_family="L1"
52346..52617
repeat_region /rpl_family="L2"
53530..53581
repeat_region /rpl_family="L2"
54254..54761
repeat_region /rpl_family="L1"
54759..55385
repeat_region /rpl_family="L1"
55659..55857
repeat_region /rpl_family="Alu"
58134..58298
repeat_region /rpl_family="MIR"
59008..59181

```

Query Match 43.9%; Score 1194.2; DB 9; Length 12331;
 Best Local Similarity 99.7%; Pred. No. 2e-216;
 Matches 1196; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

QY 1502 aagaataaatacagcgttgcagcccttccctgcttgcagcagtgctcgaagaaga 1561
DB 94286 AAGCAATAAATACAGTGTGGATCCCTTTCCTGTTCTCAGTGTGCTGAAAAAGA 94227
QY 1552 ctgaacaaaagaagaatgtgataatgcaacataatcttcagaagaatgtccaggaagatg 1621
DB 94226 CTGAACAAAAGGAAAAAGTGAATGCAACATATTTCTCAGAAATGTCAGGAAGAG 94167
QY 1622 atacaacagtgagagagagaattccctgtataagaagaccaggaagaactgaaaaaagc 1681
DB 94166 ATACAACAGTGAAGAGAGCAATTTCTGTATTAAGAGAGCCAGAAACTGAAAAAAGC 94107
QY 1682 tccgtttatctcagagccacccccccacccctcaaatgaatgtgagggcttaatgaga 1741
DB 94106 TCCGTATTATTTCAAGAGCCCAATCCCCACCCCTTCAATGAATGAGAGGGCTTAATGAGA 94047
QY 1742 aatagagcaataaagtccatggttaagtacagctgagatgacataagaacagaattta 1801
DB 94046 AATATAGTAATTAAGTTCACATTTAGTACAGCTGAAGATGACATTAAGACAGAAATTTTA 93987
QY 1802 cacaactaccctacataaaacaacaagaagatgcatctctgacatttcgagacacacat 1861
DB 93986 CACACTTACCTTACATTAATAAACAACAGAGATGCATTTCTTACATTTCCGACACACAT 93927
QY 1862 taagtgaaatgacttagaagaactaaggtagatgacataataatgtaacacagggcat 1921
DB 93926 TAAAGTAAATGACTTAGAAGAACTAAGGGTAGATCACAATAATGATTAACATTAACAGGCAT 93867
QY 1922 ctgtacatgcttctgattcagtagacataaagtatggtatctcaacaaaacagaagtcag 1981
DB 93866 CTGTACATGTTCTGATTTCACTACAGTAATATGTGATCTTAACCAAAAACGAAGTCTAG 93807
QY 1982 ataccgctctttccacaagaagatctcaaggaagaagacctcaatcaatatttacc 2041
DB 93806 ATACGTGTGTTTTTCCAGCAAAAGATCTCAAGAAAGAAAGACCTTCAATATTTTACTC 93747
QY 2042 atgattctggtctgataaacaataacagltcaagaagacacctaactgttcagcaaaag 2101
DB 93746 AATGATTTCTGTGTATTAACAATAAAGTTCACAAGAGACCACTTAATGTTCAAGCAAAAG 93687

```

```

QY 2102 ccccatccatactcctcctgaggaaccacatgaatgtaactcaagaatagatgatt 2161
DB 93686 CTCCTTCCATCTCTCTCTGTAGGAACCAATGATGATGATGATGATGATGATGAT 93627
QY 2162 tacctctggttaaaatatacgcgaagtgaaaaataatataatagaggaatagaaaaaaga 2221
DB 93626 TACCTTCTGTGTAATAATACATCAAGAAAGTAATAATATTAGAGCAAAATAGAAAAAGA 93567
QY 2222 atctggaaccaaagctgtaattgataaagaactgtaatttttcacagaagaagaana 2281
DB 93566 AATCTGAACCAAAATGCTGTAATTTGATTAATAAGAACTGAATTTTACACAGAAAGAAAGA 93507
QY 2282 gaattgtgagtcaccggtgacagctcttactagactgtttcaagctagtagaagaat 2341
DB 93506 GAATTTGTAGTTCACCGGTACAGTCTTTTACTGACTGTTTCACTGTAAGAGAAAT 93447
QY 2342 cagaatttttggtttcaacaagctcacagaagaagaagtgataaagcaatglttagata 2401
DB 93446 CAGAAATTTTGGGTTTCAACAACCTACACAGAAAGAGTGTAATAGCAATGTTTAGATA 93387
QY 2402 ttggggaagaggaataatcagataatcgttaacagcgttttctcgtccctcaact 2461
DB 93386 TTTGGGAAGAGGAAATTCAGATTAATCTGTAAACAGCGTTTTCCTCCCTCAACTT 93327
QY 2462 ctacatttactgctttagaatttaaaatgcacttttcagaagagataagaatca 2521
DB 93326 CTACATTTACTGCTTTTGAATTTAAAAATGCACTTTTCAGAGATGATTAAGATCA 93267
QY 2522 taltcttgaatttlatataatagatgaaatcttcttgaatttttaccagcttctg 2581
DB 93266 TATTTCTGAAATTTTATTAATAATGATGGAATTTCTTGAATTTTTCACAGCTTTGT 93207
QY 2582 ttacagaaccaaagttaaatatlaaaataaataatltgcaattttctacagaatgata 2641
DB 93206 TTACAGACCCCAATGTAAATATTAATAATAATTTTGCATTTTCTACAGATTTGAATA 93147
QY 2642 cctgttaagaaaatttaagaataactgtgacgtgctgttttaccataaaaaa 2700
DB 93146 CCTGTAAAGAAAAATTAACAGAAATAACTTGTGACTGCTGTGTTTACATTTATATA 93088

```

```

RESULT 8
LOCUS AC024947 163031 bp DNA PRI 24-AUG-2000
DEFINITION Homo sapiens chromosome 01 clone RP11-200110, complete sequence.
ACCESSION AC024947
VERSION AC024947.4 GI:9887599
KEYWORDS HTG.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 163031)
AUTHORS Smith,D.R.
TITLE Genome Therapeutics Corporation Sequencing Center: Human Genome
Sequence Data
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 163031)
AUTHORS Smith,D.R.
TITLE Direct Submission
JOURNAL Submitted (03-MAR-2000) Genome Therapeutics Corporation, 100 Beaver
Street, Waltham, MA 02453, USA
REFERENCE 3 (bases 1 to 163031)
AUTHORS Smith,D.R.
TITLE Direct Submission
JOURNAL Submitted (24-AUG-2000) Genome Therapeutics Corporation, 100 Beaver
Street, Waltham, MA 02453, USA
COMMENT On Aug 24, 2000 this sequence version replaced gi:8569066.
FEATURES
source
1..163031
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="01"

```

```

/clone="RP11-200110"
/clone.lib="RP11-11"
BASE COUNT 48079 a 30305 c 32059 g 52588 t
ORIGIN

```

```

Query Match 43.9% Score 1194.2; DB 9; Length 163031;
Best Local Similarity 99.7%; Pred No. 2,1e-216;
Matches 1196; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

```

OY 1502 aaagaataaatacagtggttgatcccttcctcgttccgagtgctgcaaaaga 1561
DB 35370 AAGAAATTAATACAGTGTGGATCCCTTTCCTGTCGCAAGTCTCGAAAAAAG 35429
OY 1562 ctgaacaaaggaagtggaattgcaacatattctcagaagaattgcagagaatg 1621
DB 35430 CTGAACAAAGGAAGAGTGGATGGCAATGTCATTTCTCAGAAAAGATTGCCAGAGATG 35489
OY 1622 atacaacagtgagagcagaattccctgtataaagagaccgagaaactgaaaaaac 1681
DB 35490 ATACAAACAGTGAAGAGCAGAAATTCCTGTATTAAGAGCCAGAGAACTGAAAAAAGC 35549
OY 1682 tccctgtattcagaagccatccccaccctcgaatggaattgagaggctaatgaga 1741
DB 35550 TCCCTGTATTTCAGAGGCCATCCGCCCTTCAAAATGAGAGGCTTAATGAGA 35609
OY 1742 aaatgagtaataacgttccatgttaagtcagctggaagatgacataagaacagaattta 1801
DB 35610 AAATGAGTAATTAATGTTCCATGTTAAGTACAGCTGAAGATGACATTAAGACGAATTTTA 35669
OY 1802 cagaactacccttacaataaacaacagaatgcatcttcacacattccgaaacacat 1861
DB 35670 CACAGCTACCTCTACATTAACAAACAGAGATGCAATCTTCATTTCCGACACACAT 35729
OY 1862 taagtgaataactagaagaactaaggtagatgacataataatgtaacatacagaact 1921
DB 35730 TAAGTGAATAACTAGAACTAAGGATAGGATACATTAATGATTAAGTAACTAGAGCAT 35789
OY 1922 ctgtacatgttccatcattcagtcacataatagtgactcaacaaacagaagtcag 1981
DB 35790 CTGTACATGTTCTCATATTTCAATGATGATGATGATGATGATGATGATGATGATGATG 35849
OY 1982 atactgctcttccagaaaggaattcagaagaagacccaatcattcaattatctc 2041
DB 35850 ATACTGCTCTTCTCCAGCAAGAGATCTCAAGAGAAAGACCTTCATTTCAATTTTCTC 35909
OY 2042 atgattcgtgctgataacaaataacagttcacagaagcactaacgtgttcaggcaag 2101
DB 35910 ATGATTCTGTGCTGATTAACAAATTAACAGTTTCACANAGACCTTAACCTTTCAGCAAGG 35969
OY 2102 ctccatctcaactcctcctcaggaacccaatgtaattgacttcaagaata tgaatg 2161
DB 35970 CTCATCTCCATCTCCTCCTGAGAACCCCAATGATGATGATGATGATGATGATGATGATG 36029
OY 2162 taaccttctgttaataaatacagtaagtaaaataataataagacgaataatgaagaaga 2221
DB 36030 TACCTTCTGTTAAATACATGCAAAAGTGAATAATATATTAGAGCAATGAAAGAAA 36089
OY 2222 atctggaacaaatgctgaattgtataaagaactgaatttttttttttttttttttttt 2281
DB 36090 ATCTGGACCAAAATGCTGAATTTGATTAAGAACTGATTTATTTACACAGAGAGAAAACA 36149
OY 2282 gaattgtagttcacgcgttacagcttctaactagactgtttcagactgctgaagaaga 2341
DB 36150 GAATTTTGTAGTTACCGGTACAGTCTTTTACTAGACTGTTTTCAGACTGATGAAGGAAT 36209
OY 2342 cagaatttttgggtttcacaaactacacagaanaagtggtatataatgatttttagata 2401
DB 36210 CAGAAATTTTGGGTTTTCACAACTACACAGAAAAGATGATATATGCAATGTTTATAGATA 36269
OY 2402 ttggaagaagaagaatacagaatactgttaacagaagtttttctcgtcccttcaact 2461
DB 36270 TTGGAAGAGGAATATCATATATGTTTAAACAGCGTTTCTGTCCTTCAACTT 36329

```

```

OY 2462 ctacatttcctggtctttagaattaaaaatgcatcttctcagaagatgaagaatca 2521
DB 36330 CTACATTTCCTGCTTTAGATTAATAAATGCAATCTTTCAGAGTGAATTAAGATCA 36389
OY 2522 tattctggaattttataaataatgataatggaatctcttaggatttttttaccagcttctg 2581
DB 36390 TATCTTGAATTTTATTAATAATGATGAATTTCTTAGAATTTTTCACACATTTGT 36449
OY 2582 ttacagaccgaatgtaataatlaaaaaataatttgcaattttctacagaattgaata 2641
DB 36450 TTACAGACCCAAATGTAATATTAATAAATTTGCAATTTCTACAGAAATGATA 36509
OY 2642 cctgttaagaagaataatcaagaataaactctgactgctctgttttaccataaaaaa 2700
DB 36510 CCTGTAAAGAAAATTAACAGAAATTAACCTGTGACTGCTGTTTTCATTATATATA 36568

```

```

RESULT 9
AC004958 163889 bp DNA HTG 12-JUN-1998
LOCUS Homo sapiens clone DJ1096F2, *** SEQUENCING IN PROGRESS ***, 6
DEFINITION
AC004958
AC004958.1 GI:3213039
VERSION
HTG; HTGS_PHASE1.
KEYWORDS
human.
SOURCE
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 163889)
Waterson,R.H.
The sequence of Homo sapiens clone
unpublished
2 (bases 1 to 163889)
Waterson,R.H.
Direct Submission
Submitted (12-JUN-1998) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA

```

```

COMMENT
* NOTE: This is a 'working draft' sequence. It currently
* consists of 6 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1
1462: contig of 1462 bp in length
1481: gap of unknown length
1481: contig of 1584 bp in length
3063: gap of unknown length
3063: contig of 1628 bp in length
4711: gap of unknown length
4728: gap of unknown length
6531: contig of 1802 bp in length
6547: gap of unknown length
6548: contig of 9882 bp in length
16430: gap of unknown length
16447: contig of 147443 bp in length.
*
* Location/Qualifiers
1.163889
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="DJ1096F2"

```

```

BASE COUNT 49499 a 30566 c 31570 g 52166 t 88 others
ORIGIN

```

```

Query Match 43.9% Score 1194.2; DB 2; Length 163889;
Best Local Similarity 99.7%; Pred No. 2,1e-216;
Matches 1196; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
OY 1502 aaagaataaatacagtggttgatcccttcctcgttccgagtgctgcaaaaga 1561

```

```

Db 74946 AAGAGTAAATFACAGTGTGGATCCCTTCTCCTGTTCTCCAGATGTCCTGAAAGAA 75005
|
|
|
Oy 1562 ctgaacaaagaagaaatggaatgtcaacatatttccagaagaatgtccagaagaatg 1621
|
|
|
Db 75006 CTGAACAAAGAAAGAAATGTGAATGCAACATATTCTCAGAAAGATGTCAGGAAGATG 75085
|
|
|
Oy 1622 atacaacgtgaagagagagaatctccgtatataagaagaccagaagaactgaaaaaagc 1681
|
|
|
Db 75066 ATACAACAGTAGAGAGAGAGATTTCCGTATTAAGAGACCAGAGAACTGAAAAAAGC 75125
|
|
|
Oy 1682 tccgttatttccagaagccatccccccatcaatgaaatgtgagaggttaattgaga 1741
|
|
|
Db 75126 TCCTGTTATTTCAGAGCCCATCCCCACCTTCATAATGAATGTGAGAGGCTTAATGACA 75185
|
|
|
Oy 1742 aaatgaagataaattgtccatglttaaglacagctgaagatgacataagacagaattta 1801
|
|
|
Db 75186 AAATGAGTAATTAATGTTCATGTTAAGTACAGCTGAAGATGACATTAAGAGCAATTTTA 75245
|
|
|
Oy 1802 cagagctacccctacataaacaacaaaggaatgcattcttgaacattccgaacacacat 1861
|
|
|
Db 75246 CACAGCTACCTCTACATATAAAACAAACAGGAATGCATTCTTGACATTTCCGAAACACAT 75305
|
|
|
Oy 1862 taagtgaaatgagacttagaagaactaaggtatagatcactataaattgtacatacagagat 1921
|
|
|
Db 75306 TAAAGTAAATGACTTAAAGAACTTAAGGGTAGATCACTATTAATGTATACATACAGCAT 75365
|
|
|
Oy 1922 ctgtacatgttctgtatttccaglacagataatagtgatgtctcaacaaacagaagtcag 1981
|
|
|
Db 75366 CTGTACATGTTTCTGTATTCAGTACAGATTAATGATGATCTCAACCAACGAAGTCTCG 75425
|
|
|
Oy 1982 atactgtcttcttccagaacaaagatctcaaggaagaagagaccttcaataattactc 2041
|
|
|
Db 75426 ATACGTCTCTTTTTCACAAAGAGATCTCAAGGAAGAGACCTTCATCAATATTTTAC 75485
|
|
|
Oy 2042 atgattctggtctgtatacaataacagttcaagaagacaccttaacttccagaagaag 2101
|
|
|
Db 75486 ATGATTTCTGGTCTGTATTAACATTAACAGTTCACAAAGACACCTTAACGTTCAGCAAG 75545
|
|
|
Oy 2102 ctccatctccatctctctctctgaggaacccaatgaattgtacttcaagaataatgagat 2161
|
|
|
Db 75546 CTCATTTCCATATCTCTCTCTGAGGAACCCCAATGATGTGACTTCAAGATATGATGATG 75605
|
|
|
Oy 2162 taaccttcgtgtataacatcatcagaaagtgaataataatattagacgaagaatagaaga 2221
|
|
|
Db 75606 TACCTTCTGGTAAATATCATGAAAAAGTAAATATATAGAGCAATATGAAAAAGAA 75665
|
|
|
Oy 2222 atctggaaccaaagctgaattgtataaagaagaaatttaatacaagaagaagaaca 2281
|
|
|
Db 75666 ATCTGGAACCAAAATCTGAAATTTGATTAAGAACTGAATTTATTTACACAGAAAGAAACA 75725
|
|
|
Oy 2282 gaatttgatgtacacgggtacagcttcttactagactgttccagaactagtgaagaaga 2341
|
|
|
Db 75726 GAATTTGTAGTTTACACGGGTACAGCTTACTTACTGACTTGTTCAGACTAGTGAAGAAAT 75785
|
|
|
Oy 2342 cagaatttttgggtttccacaagctcacagaagaagatggtatagcaatgttttagata 2401
|
|
|
Db 75786 CAGAAATTTTGGGTTTACAAACCTACACAGAAAGAGTGTATAGCAATGTTTAGATA 75845
|
|
|
Oy 2402 ttgttggaagagaagaattcagataatctgtttacagagcttttccgttcccttaact 2461
|
|
|
Db 75846 TTTGGGAAGAGAAAAATTCAGATTAATCTGTATACAGCGTTTTCGTCCCTTCAACTT 75905
|
|
|
Oy 2462 ctacattacttgcgtctttagaatttaaaaaatgcatcttccagaagtgaataagatca 2521
|
|
|
Db 75906 CTACATTTTACTGCGTTTGTAGATTTAAAAAATGCACTTTTCAAGAAGATGAAGATCA 75965
|
|
|
Oy 2522 tatctttaaattttataataatgataagaaatctctagagtttttttaccagctttgt 2581
|
|
|
Db 75966 TATTTCTGAAATTTTATTAATAATATGATGAAATCTTGAAGATTTTTCACAGCTTGT 76025
|
|
|
Oy 2582 ttaacagaccaaatglaaataataaataatcttgaatttctacagattgaaata 2641
|
|
|

```

```

Db 76026 TTACAGACCCAAATGTAAATATTAAATAATATTGCAATTTCTACAGAAATTGAATA 76085
|
|
|
Oy 2642 cctgttaaagaanaattacagaataaactgtgactgtgtctgttttaccatataaaaaa 2700
|
|
|
Db 76086 CCTGTAAAGAAAAATTTACAGAAATTAACCTTGACTGCTGCTGTTTACATATATATA 76144
|
|
|
RESULT 10
MIMBDF4
LOCUS MIMBDF4 2309 bp mRNA ROD 11-NOV-1999
DEFINITION Mus musculus mRNA for DBF4-related protein, complete CDS.
ACCESSION AJ003132
VERSION AJ003132.1 GI:6006562
KEYWORDS DBF4-related protein.
SOURCE house mouse.
ORGANISM Mus musculus
MIMBDF4: Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 2309)
Lepeke, M., Puterbaugh, V., Stalib, C., Kneissel, M., Berger, C., Hoeft, K.,
Nanda, I., Schmid, M., and Grunmt, F.
Identification, characterization and chromosomal localization of
the cognate human and murine DBF4 genes
Mol. Gen. Genet. 262 (2), 220-229 (1999)
99444905
2 (bases 1 to 2309)
Lepeke, M.
Direct Submission
Submitted (17-DEC-1997) Institute of Biochemistry, Am Hubland,
Muerzberg D-97074, Germany
JOURNAL location/qualifiers
FEATURES
source 1..2309
/organism="Mus musculus"
/db_xref="taxon:10090"
95..2086
/gene="MIMBDF4"
95..2086
/gene="MIMBDF4"
/function="Involved in initiation of DNA replication"
/codon_start=1
/product="DBF4-related protein"
/protein_id="CAB56847.1"
/db_xref="GI:6006563"
/translation="MNELETKRIHSAKAPLPDGIQDRBNKRNRSLSKSKADNLEMSK
PLWGEIFLIDLPSTICEKIDKIDKELGAREEPLSKDISEVSKKAKAKAQTG
SPVSPESAVYAEETSPSPSHDGSFQDVCUSRGKLEKAKVDHDFIPANSILS
NALSMGVILHIDIRYIEOKKRLSALKSSASGAGKAGKIOKTRGLRKP
FLKVEDVNRQYRPFLQIPSLPCINYEFLQPCSPDIDKSSVQOAPKLRIMNDG
KCGTRVOLQLEKRRKRGCECCLOKEDLETHLSERKRNRAQSNQYVNDYQSLY
FDEYVGRDTPQKRIKIRISVGSLSVSNVANKTAPKPLDEPRFQDVESSGHL
KPNQOYETQKPEEKHGFASPTYSAGLGCDBKPVSMFNASBPDEQEAQPLR
DSTPEHOVTEBRNDQEVDPAPVQSGQGVSHLSTVSNLPQDLADITOLSKDL
OEKGFHVYIGHASDLVALNTSEKQLETKMARTPCSPQPHEDQENMNLPCGIQK
VRMLIGQOKNAPEASDELDRKRTYELPAHEDRTGSPVQSLDLFQTSSEKSEFLGFT
GYTENSQICDVLDIWEENSSSTILSTFSSSTSAFVGF"
BASE COUNT 740 a 466 c 526 g 577 t
ORIGIN
Query Match 41.5%; Score 1129.6; DB 10; Length 2309;
Best Local Similarity 71.6%; Pred No. 3e-204;
Matches 1668; Conservative 0; Mismatches 554; Indels 106; Gaps 10;
Oy 445 ggaagcgtgcttccgagctgcccgtgagacatttcccgagaccagcatgtagtg 504
|
|
|
Db 22 ggaagcgttctgttcggcgccgcttcacatcgccgctgagacaccagatgtagtg 81
|
|
|
Oy 505 ccggcgagatgcacatgaactccggagccatgagatccacagataaagacatttccagg 564
|
|
|
Db 82 CCTGTGGGTGGCCATGAACTCGAGACCATGAGGATCCACACAAAGCACCCTCCCGG 141
|
|
|
Oy 565 tgaatccaatgcaaaaatgaaataaagacacatctcgaactctcgaactcgaactgataa 624
|
|
|

```

Db 142 TGGCATCCACAGAAATGAAAGATCGACATCCTTGAAATCCCTGAGCGGATMA 201
Qy 625 caagcagagaaatccaaatgtaagccacttgggaaagatlttacccttgacttacc 684
Db 202 TGGAGTGGAAAAATCCAAATATATAGCCCTTGGGAAAAATTTTACCTTGACTGCG 261
Qy 685 ttcctgcaacatctgtaaaacttcaaaagacattgaagatctgggagggcgttga 744
Db 262 TCTATACCACTATGTAGAAACTACAAAAGATATMAAGAACTCGAGAGGGATGTA 321
Qy 745 agaattctcagcaagatactcagttatcttcaatlaagaaggaagctaaattgc 804
Db 322 AGAATTTCTCAGAAAGATATCAGTTACTTGTTCAAATMAAAGCAAGCAAGATATGC 381
Qy 805 agaaacctgggtcgaattctcctgtaccagtcagaatctgcatatctgcagaac 864
Db 382 ACAGACATTTGGGTGGCTTCTCTGTACCAAGTCCAGAAATCTCATATCTGCAAAAC 441
Qy 865 caattcaactatccagccatgatagtaagttcatcttaagtcaccagacagtggtt 924
Db 442 CACCTCTCCTCCAGCCACATCGAAAGTTCCTTTAAGTCTCAACAGAGATATGTTT 501
Qy 925 aagcagaagaaatbatagttgaaaaagctatcaaggaaccaatgtatcttccctcaa 984
Db 502 GAGCCGAGAAAAATTTGAGCTGAAAAAGCTGTGAAGATCATATATTTATCTCTGTA 561
Qy 985 tagtataatcaaatgctctgtcattgaggagtaaaatcttccatctgagacattag 1044
Db 562 TAGTATATATATCAATGCTTATCATGGGAGTAAGATTTCTGATTTATGATACATTCG 621
Qy 1045 ataccacattgcaaaaagaaagaaagatgtaatttactcaagaatcaagttcagtl 1104
Db 622 ATACTGATTTGAAAGAAAGAAAGAGAGAGAGTGTGACTCAAGAAATCCAGTCTCTG 681
Qy 1105 aagaagctggggcaaaagatgttggtagtggcacaacaaagaaagaaagaaactcaa 1164
Db 682 GAAGGATGGGGGAAAAAGCTGGCCCTGGAAATACAAAAGACAGAGAGAGACTCAA 741
Qy 1165 aagaccttttgaagaagtgtaagatgatgagc----- 1196
Db 742 GAAGCCTTTTAAAGGTTGAAGATGTGAACCGTGCTACAGGCCATTTTACCTTACGCT 801
Qy 1197 -----aagacctggcgatcattgtagtga 1223
Db 802 CCCCAGTCCCGGTATTAACCTATTCTTCAGAAAGCCCTGCATCGCTTGATATAGA 861
Qy 1224 caagccactagtatgcaaaagcaactcaaggttaactaaagaatccaaagatggcga 1283
Db 862 GAAGTATATCCAGCTGCAGAAAGAGCTCAGCCCAAGCTCAGATTAATTAATGACGCGCA 921
Qy 1284 taagtaatgtgaacatcaatcaactcaactcaagttgaaagaaagaaagaaagatgtg 1343
Db 922 TAAAT---GTGAAACCCCAAGTTCAACTCAGTTGAAGAAAGAAAGAAAGAAAGATTTTG 978
Qy 1344 tgaatgtctgtcgcaaaata tgaagatctagaactccacttcaagtgagcaacag 1403
Db 979 TGAATGTCTTACGAAATATGAGAAATCTCGAAACTCAGCTTCTGAGTGAAGAAACAG 1038
Qy 1404 aaacttgcacagagtaaccagttatcaagttgtgtagtattgtatcttaagttagttt 1463
Db 1039 GAACCTTGGACAGAGTAACCAAGTATCAAGTTGTGATGATATGATCTCAGTTAGTTT 1098
Qy 1464 tgaatttggtaagatgaaaggaacacaccttaaaagaaagaaagaaataaatacatgttgg 1523
Db 1099 CGATTTTGTGAAATATGAAAGACACACCTCAAAAGAAAGAAATATGATGATGACGTGCG 1158
Qy 1524 atcccttccctgttctgcgaagtgctgcgaagaaagaaagaaagaaagaaagaaag 1583
Db 1159 GTCCCTTCTCTCTGTTCTGCAAAATGTCTCTGAAACACTGCAACGAGAGAAAGCAACT 1218
Qy 1584 atgcaacatattctcagaagaatgcaggaagaatgatatacaacagtgaaagagcaga 1643
Db 1219 CTGTGAGCCCAATTTTCCAGAAAGAGCTGGGCAAGCAGTGGGACCTGCTCAAGCCGAA 1278

Qy 1644 tttcctgtataagaagaccagaaactgnaaaaaagctcctcgttatttcaagagccat 1703
Db 1279 CTCTCATGTATGAAGAAACCCAGAAACCTGAAAGAAAGCATATGGGTTTGCTTCAGGCCAC 1338
Qy 1704 cccccccttcaaatggaattgagagggcttaatggaanaatgagtaaatgttccat 1763
Db 1339 AACCTATTCTTCAGCTGATGTAAGAAAGATGTA-----CAGCAAAACCATGAGACAT 1389
Qy 1764 gtttaagtaagctgtaagatgataaagaacgaatttatacagctaccctcactataaaa 1823
Db 1390 GTTCATATGCAAGTGAACCTTGACCAAGACAGAGATGTGACACAGCTTCTGCGTGACAG 1449
Qy 1824 caaaggaatgcaattctgcaattccgaaacacataaagtgaanaatgactagaaga 1883
Db 1450 CACACCGGAAC-----ACAAAGTGAACCGAGGTTAGAAATGAC---GGAGA 1491
Qy 1884 actaaggttagatcactataaatgtaacataacaggaatctgtacatgttcttgaattcag 1943
Db 1492 ACAGAGAGTGAACCTGCTCCAGCGGTGCGCATGTGAGCAAGTTTCTCATCTCAG 1551
Qy 1944 tacagataatgtagtgcacaaacaaacagaagtcagaatctgtgcttcttccagaaa 2003
Db 1552 CACAGAAAGTATCTACCCCAACCAAGTTAGTGCAGACATTTACACAGCTTTGACGAAA 1611
Qy 2004 ggaatcgaagaaagagacctcatcaataattactcaatgtatctgtctgataaacaat 2063
Db 1612 AGATCTCGAAGAAAGGCTTTCATGTTGTGATGATGATGATGATGATGATGATGATGAT 1671
Qy 2064 aaacagttcaagagcaactactgtctcaggaagagctcactcactcactcctcga 2123
Db 1672 AATATCATCAAAAGACACACTTAAATCAAGACGAGACGACACCTGATGATGATGATGAT 1728
Qy 2124 ggaacccaatgaaatgtgactcaagaatgtaagtttactccttctcgtgtaaatatctg 2183
Db 1729 GGAACCTCATGAGCTGTGACATGAGACATGAGAAATTTCCCTGCTGCTGCTGCTGCTGCT 1788
Qy 2184 aaaaatgaaataatattgagaagaaatagaagaataatcgtgaaccaaagtcgaat 2243
Db 1789 GAAAGTATGAGATCTATAGGGCAACA---GAAAGCAAAATGCTGAACGAGAGCTGAAAT 1845
Qy 2244 tgataaa---agaactgaaatttatlacacaagaagaacagaatttgaatcagcgt 2300
Db 1846 GGAATTAAGAGAACTGATATCTTCGACATGAAAGACAGAACTTGTGGCTCACAGT 1905
Qy 2301 aacagcttactagactgttccagaactagtgaaagaaatcagaatttgggtttcac 2360
Db 1906 ACAGTCTCTATTTGACCTTTTTCAGACAAAGAGAGAGAAATTCAGAAATTTTGGGTTTAC 1965
Qy 2361 aagctacagaaagagtgatatagtcaatgtttagatatttgggaagaaagaaatc 2420
Db 1966 AGCTTACATGAAACAGTGAATATGATGATGATGATGATGATGATGATGATGATGATGAT 2025
Qy 2421 agataatgtttaaagcgttttctcgtcccttcaacttcaacttcaacttcaacttca 2480
Db 2026 AAGTACCTGTTCACAGCTTTTTCCTCCCTTCAACTTCGATTCGATTCGATTCGATTCGAT 2085
Qy 2481 gaatttaaaaaatgcaacttccgaagtgatagatataatcttgaatcttata 2540
Db 2086 GTATTAAATGTCTTTTCAGACAGCTATATTCCTCAAGTTTATTAATATCAATAT 2145
Qy 2541 aatagtatggaatctttagaatttcttaccagcgttgttaccagaccaaagttaa 2600
Db 2146 ATTAATTAAGTGTGGGGTTTGTGTTTTCGCAACTTTTATGAGAGCTGTATTA 2205
Qy 2601 taatata--aaataaataatcttcaatcttcaagaatgtaatacctgtttaaagaatc 2658
Db 2206 TATTAATTAAGCTGTGTTGTGTAATTTTACAGGATTTGACTACCTGT---ACAAAT 2261
Qy 2659 acagaataacttggactgttctgttcttatacattaaaaaanaanaa 2706
Db 2262 ATACATTAATAAATTAATGACTTAATTTATTAATAAATAAATAAATAAATAAATAAATAA 2309

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Alsbrooks, S.L., Amaralunge, H.C., Are, J.R., Banks, T., Barbara, J., Benton, J., Blinige, K., Blankenburg, K., Bonino, D., Bouck, J., Bowler, S., Briley, M., Brown, E., Brown, M., Bryant, N.P., Puhay, C., Burch, P., Burckell, C., Burrell, K.U., Byrd, N.C., Carron, T.F., Carter, M., Carrasco, S.R., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Z., Chowdhury, I., Christopoulos, C., Cleveland, C.D., Cox, C., Coyle, M.D., Dalhorne, S.R., David, R., Davila, M.L., Davis, C., Day, A.L., Ding, Y., Dinh, H.H., Douhal, K.J., Delgado, O., Denny, A.L., Ding, Y., Dinh, H.H., Douhal, K.J., Delgado, O., Dugan-Rocha, S., Durbin, K.J., Earnhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escotto, M., Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P., Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R., Gorrell, J.H., Guevara, M., Gunaratne, P., Hale, S., Hamilton, K., Harris, C., Harris, K., Hart, M., Hawlak, P., Hawes, A., Hollins, B., Homsl, F., Howard, S., Huber, J., Hulik, S., Hume, J., Jackson, L.E., Jacobson, B., Jia, Y., Johnson, R., Jollivet, S., Joudah, S., Karlsson, E., Kelly, S., Khan, U., King, L., Korah, J., Kovar, C., Kravtsov, J., Kuresh, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L., Li, J., Li, Z., Lichteberg, O., Lieu, C., Liu, J., Liu, W., Louissegh, H., Lozano, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J., Maheshwari, M., Mapua, P., Martin, R., Martindale, A., Martinez, E., Massey, E., Mawhinney, E., McLeod, M.P., Meador, M., Mei, G., Metzger, M., Miner, G., Miner, Z., Mitchell, T., Mohabati, K., Morgan, M., Morris, S., Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N., Nguyen, N., Nickerson, E., Nwokenwo, S., Oguchi, M., Okunolu, G., Oregan, N., Oyler, R., Pace, A., Payton, B., Peery, J., Perez, L., Peters, L., Pickens, R., Prims, E., Pyl, L., Quiles, M., Ren, Y., Rivera, S., Scott, G., Shen, H., Shostakov, N., Sisson, I., Sodergren, E., Sonalke, T., Sparks, A., Stanley, H., Stone, H., Sutton, A., Svatek, A., Taber, P., Tamerisa, A., Tamerisa, K., Tang, H., Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S., Usmani, K., Vasquez, L., Vera, V., Villalón, D., Vinson, R., Wall, R., Wang, S., Ward-Moore, S., Warren, R., Washington, C., Wallington, S., Williams, G., Williams, A., Wiczyski, R., Wooden, S., Worley, K., Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorilla, S., Nelson, D. and Gibbs, R.

Unpublished
Direct Submission
2 (bases 1 to 232903)
Morley, K.C.
Submitted (18-APR-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
Project Information
Center project name: HNCU
Center clone name: CH230-1C19
Summary Statistics
Sequencing vector: Plasmid: M7789
Chemistry: Dye-terminator Big Dye 3.1 of reads
Assembly program: Phrap, version 0.990329
Consensus quality: 230627 bases at least Q40
Consensus quality: 245439 bases at least Q30
Consensus quality: 253180 bases at least Q20
Estimated insert size: 222874; sum-of-contigs estimation
Quality coverage: 5x in Q20 bases; agarose-gel estimation
Quality coverage: 5x in Q20 bases; sum-of-contigs estimation

NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/genbank_dirft_data.html).
NOTE: This is a 'working draft' sequence. It currently
consists of 35 contigs. The true order of the pieces
is not known and their order in this sequence record is
arbitrary. Gaps between the contigs are represented as

runs of N, but the exact sizes of the gaps are unknown.
This record will be updated with the finished sequence
as soon as it is available and the accession number will
be preserved

28075	contig of 28075 bp in length
28076	gap of unknown length
28175	contig of 1438 bp in length
42563	gap of unknown length
4263	contig of 1976 bp in length
4264	gap of unknown length
62433	contig of 1668 bp in length
62433	gap of unknown length
79501	contig of 1335 bp in length
92955	gap of unknown length
92956	contig of 10802 bp in length
103758	gap of unknown length
103858	contig of 13202 bp in length
117059	gap of unknown length
117160	contig of 8843 bp in length
126003	gap of unknown length
126003	contig of 8306 bp in length
134409	gap of unknown length
134509	contig of 9553 bp in length
144062	gap of unknown length
144062	contig of 9144 bp in length
153305	gap of unknown length
153305	contig of 7994 bp in length
153406	gap of unknown length
161400	contig of 6779 bp in length
161500	gap of unknown length
168278	contig of 7097 bp in length
168379	gap of unknown length
168379	contig of 3377 bp in length
175876	gap of unknown length
175876	contig of 4724 bp in length
180953	gap of unknown length
181053	contig of 3578 bp in length
185777	gap of unknown length
185777	contig of 3931 bp in length
189544	gap of unknown length
189544	contig of 4713 bp in length
193485	gap of unknown length
193485	contig of 3072 bp in length
198298	gap of unknown length
198298	contig of 3728 bp in length
201471	gap of unknown length
201471	contig of 3165 bp in length
205299	gap of unknown length
205299	contig of 2688 bp in length
208584	gap of unknown length
208584	contig of 4012 bp in length
211372	gap of unknown length
211372	contig of 1710 bp in length
215484	gap of unknown length
215484	contig of 1501 bp in length
217294	gap of unknown length
217294	contig of 1796 bp in length
218995	gap of unknown length
218995	contig of 2263 bp in length
220791	gap of unknown length
220791	contig of 1257 bp in length
223153	gap of unknown length
223153	contig of 1349 bp in length
224511	gap of unknown length
224511	contig of 1829 bp in length
225960	gap of unknown length
225960	contig of 1046 bp in length
227889	gap of unknown length
227889	contig of 1348 bp in length
229035	gap of unknown length
229035	contig of 1179 bp in length
230483	gap of unknown length
230483	contig of 1042 bp in length
231762	gap of unknown length
231762	contig of 1042 bp in length


```

Db 37578 GGGGAAATTCAGAAAGTTGGGTTTACACCTACAGTAAACAGTGAATCTGTGATG 37519
QY 2393 tttaagatatttggaagaggaattcagaataatcgttaacagc-gttttctgcgtc 2451
Db 37518 -----AAATTCAGTACCGCTGTGTCACACATTTTCTTCC 37481
QY 2452 ccttaacttctacattactggtccttaagaataataaataatcattcttcaagaatg 2511
Db 37480 CCGTCACCTTCCTCATTCTCGATTTCGATTGATTAATG-----TGCCTTTCACAACTG 37427
QY 2512 ataaggaatcattcttgaatttctataatcgtat-----ggaatctctagat 2564
Db 37426 ACA--GCTATATTCCTCGAGGTTTATTAATACATATATACAGAGTGGGAGGCT 37370
QY 2565 tttttaccagcttctttacagaccacaatgtaataatcattataaataat-ttgcac 2623
Db 37369 TGTTTTGGCCAAATTTGTTATGGAACCGTGTAAATATTAATTAAGACAGATTGTAAT 37310
QY 2624 ttctacagaattgtaattactgttaagaataataataacttctgactgctt 2683
Db 37309 TTTTACAGATCGTGTACTGT-----ACAAATTATACAAATAAATTTGCACTTAT 37254
QY 2684 gtttaccattaaaaaataaaaaa 2710
Db 37253 TAAAAAATATACACCAACTAATA 37227

```

```

RESULT 13
AC003083/C AC003083 136823 bp DNA PRI 21-DEC-1999
LOCUS Homo sapiens BAC clone CTB-60N22 from 7q21, complete sequence.
AC003083
AC003083.1 GI:3947435
VERSION
KEYWORDS hng.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE
AUTHORS Kellen,J., Smith,A. and Gibson,A.
TITLE The sequence of Homo sapiens BAC clone CTB-60N22
JOURNAL Unpublished
REFERENCE
AUTHORS 2 (bases 1 to 136823)
TITLE Direct Submission
JOURNAL Submitted (06-NOV-1997) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
REFERENCE
AUTHORS 3 (bases 1 to 136823)
TITLE Direct Submission
JOURNAL Submitted (15-JAN-1998) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
REFERENCE
AUTHORS 4 (bases 1 to 136823)
TITLE Direct Submission
JOURNAL Submitted (02-DEC-1998) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
REFERENCE
AUTHORS 5 (bases 1 to 136823)
TITLE Direct Submission
JOURNAL Submitted (21-DEC-1999) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
COMMENT
----- Genome Center
Center: Washington University Genome Sequencing Center
Center Code: WUGSC
Web site: http://genome.wustl.edu/gsc
Contact: saplens@wustl.wustl.edu
----- Summary Statistics
----- Center project name: H_K060N22

```

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
 All regions were double stranded or sequenced with an alternate chemistry; an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:
 The sequence of this clone was established as part of a mapping and sequencing collaboration between the NHGRI Chromosome 7 Mapping Project (Eric D. Green, Director), John D. McPherson in the Department of Genetics (Washington University), and the Washington University Genome Sequencing Center. For additional information about the map position of this sequence, see <http://www.nhgri.nih.gov/DIR/CTB/CHR7/>, send <mailto:egreen@nhgri.nih.gov>, or see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:
 Clone CTB-60N22 is from the first release of the human BAC library CTB-978SK-B. The library contains cloned DNA from the male fibroblast cell line 978SK. See: Shizuya et al., Proc. Natl. Acad. Sci. USA 89:8947 (1992); U-J. Kim et al., Genomics 34:213-8 (1996). This clone is available from Research Genetics, Inc. (<http://www.resgen.com>).
 VECTOR: pBelBAC11
 Selection: chloramphenicol

NEIGHBORING SEQUENCE INFORMATION:
 The clone sequenced to the left is CTB-135C18, 200 bp overlap. The clone sequenced to the right is CTB-60P12, 200 base overlap. Actual start of this clone is at base position 81036 of CTB-135C18; actual end is at 36622 of CTB-60P12.

FEATURES
 source
 1..136823
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="7"
 /map="7q21"
 /clone="CTB-60N22"
 /clone_1lb="CTB-978SK-B"
 /note="CpG island (GC=66.5, o/e=0.90, #CpGs=140)"
 misc_feature
 repeat_region
 2977..3119
 /note="CpG island (GC=66.5, o/e=0.90, #CpGs=140)"
 repeat_region
 3498..3681
 /rpt_family="L2"
 /rpt_family="MIR"
 repeat_region
 4537..4608
 /rpt_family="MIR"
 repeat_region
 4609..4908
 /rpt_family="Alu"
 repeat_region
 4909..5002
 /rpt_family="MIR"
 repeat_region
 5495..5658
 /rpt_family="Alu"
 repeat_region
 5659..5737
 /rpt_family="L1"
 repeat_region
 6015..6103
 /rpt_family="L1"
 repeat_region
 6210..6304
 /rpt_family="L1"
 repeat_region
 6318..6551
 /rpt_family="L1"
 repeat_region
 6552..6764
 /rpt_family="L1"
 repeat_region
 6768..7030
 /rpt_family="MIR"
 repeat_region
 7272..7610
 /rpt_family="Retroviral"

```
repeat_region      7403..7628
                    /rpl_family="MER4-group"
repeat_region      7660..7898
                    /rpl_family="retroviral"
repeat_region      8013..8755
                    /rpl_family="L1"
repeat_region      8777..9094
                    /rpl_family="Alu"
repeat_region      9389..10058
                    /rpl_family="MER1-g"
repeat_region      10069..10231
                    /rpl_family="L1"
repeat_region      10626..11262
                    /rpl_family="L1"
repeat_region      11263..11827
                    /rpl_family="retroviral"
repeat_region      11828..12240
                    /rpl_family="L1"
repeat_region      12241..12532
                    /rpl_family="Alu"
repeat_region      12533..12548
                    /rpl_family="L1"
repeat_region      12549..12849
                    /rpl_family="Alu"
repeat_region      12850..13189
                    /rpl_family="L1"
repeat_region      13205..13372
                    /rpl_family="Alu"
repeat_region      13376..13881
                    /rpl_family="L1"
repeat_region      13882..14120
                    /rpl_family="L1"
repeat_region      14084..14283
                    /rpl_family="L1"
repeat_region      14286..14760
                    /rpl_family="L1"
repeat_region      14765..15133
                    /rpl_family="MALR"
repeat_region      15316..15669
                    /rpl_family="MALR"
repeat_region      17202..17287
                    /rpl_family="MALR"
repeat_region      17818..18025
                    /rpl_family="L1"
repeat_region      18030..18105
                    /rpl_family="L1"
repeat_region      18115..18424
                    /rpl_family="Alu"
repeat_region      18791..18835
                    /rpl_family="Alu"
repeat_region      18888..19165
                    /rpl_family="(CATR)n"
repeat_region      19778..20821
                    /rpl_family="Alu"
repeat_region      20875..20991
                    /rpl_family="Retroviral"
repeat_region      21983..42089
                    /rpl_family="MIR"
gene
CDS
    /gene="WUSC:H_RG060N22.1"
    join(<21983..22046,24028..24134,28390..28457,30361..30485,
    31216..31389,34475..34584,36624..36705,41528..41608,
    41977..42089)
    /gene="WUSC:H_RG060N22.1"
    /note="mitochondrial carrier protein-like: similar to
    009461 (PID:g2497990); H_RG060N22.1"
    /codon_start=1
    /evidence=not_experimental
    /protein_id="AAC82534.1"
    /db_xref="GI:2588623"
    /translation="IVPVIDVVKIRLQAQNNPDKRGKGFVYNSGLMDLHCCEBGNK
    LMKPGMFGGTLDIAFKIRNKGKSLMSCLPTLVAVPATVLYIFCYDQLSLR
    SKLENETCIPIVAGIARFGAVVISPLEIRIKMQSKFSYELAHFVSKVSEGD
    WISLRGAPVILRDVPSAMVWYNYELIKMLKESGLIEPTFMINFTSGLSGSA
    AVATLPDVVTKOTOLMTYESHKISMLHMSWTWIMKNIYAKNGFSGLSGLIPRL
```

```
repeat_region      23159..23457
                    /rpl_family="Alu"
repeat_region      23507..23784
                    /rpl_family="Alu"
repeat_region      24328..25740
                    /rpl_family="L1"
repeat_region      25741..26036
                    /rpl_family="Alu"
repeat_region      26037..26232
                    /rpl_family="L1"

Query Match      20.4%; Score 554.6; DB 9; Length 136823;
Best Local Similarity 99.3%; Pred. No. 3.6e-95;
Matches 557; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY      5  cggcagagctctctgaagctcgccaaagactgaagcggcgagcccgagctc 64
      1  |||||
DB      2118  CGCGCCAACTCTGAGGCTGCGCCAAAGACTGAAGCGCGGACCGAGCCGGGTCT 2059
QY      65  ggaactgaagagcagcgaatgaagcgcggttagagggcggaacacacactgcagggc 124
      1  |||||
DB      2058  GAGACTGAGAGCAACGGAATGAGGCGGGGTAGAGCGGAACAACAACCTGCAGGGCC 1999
QY      125  aagacgagcgcgagaaagacgagcgctgaaggggcgggcgcgagcgagaaagga 184
      1  |||||
DB      1998  AAGACGAGCGCGGAGAGACGCGCGCTGAGGGGGCGGCGCGCGCGAGCGCGAAGGCA 1939
QY      185  ggcagagggcgagcgcgagcgagcgagcgcgctgagcgagcgagggcgagcgcg 244
      1  |||||
DB      1938  GGCAGAGGGGAGGAGCGAGGCGAGGCGGCGCGCTGCGCTGAGACGGGGGGCGCG 1879
QY      245  cgtatcgcgcgcgcgcgcggtgagcggtttcaaatctcaacgcgcgagccactcg 304
      1  |||||
DB      1878  CGTATCGCGCGCGCGCGCGCTGAGCGCGCTTCAAAVCTTCAACCGCCCGACGCACTCG 1819
QY      305  ttgtgcttgcgccttccttcgcgcgcttggaacggatccggcccggaacccga 364
      1  |||||
DB      1818  TTGTGCTTGGCGCTTCTCTCTCGCGCTTGAGCCGGATCCGGCCCGAACCCTCA 1759
QY      365  cctgcagacgcgtaacctacttcgtaagggcgtatgcgtgcgaggaagagagcgcg 424
      1  |||||
DB      1758  CCGCAGAGCGCGGTACTCTACTGCTAGAGGCGGTAGCTGGGGAGAGAGAGCGCG 1699
QY      425  cgtctgtcaaacagcgcgaggaagcggtgcttcgcggcgctccgggtgcgacattct 484
      1  |||||
DB      1698  CGTCTGTCAACAGCGCGGGGGAAGCGCTTTCGCGGCTGCCGCTCGACACTTTCT 1639
QY      485  ccgagccagcatgtagtgcggcgagctgcataaactccggaagcattgaagatccac 544
      1  |||||
DB      1638  CCGAGCCAGCATGTAGTGCGCGGCGACTGCCATGAACTCCGAGCCATGAGATCCAC 1579
QY      545  agtaagacattccaaggt 565
      1  |||||
DB      1578  AGTAAAGACATTCCAGGCT 1558

RESULT  14
AX070167      685 bp      DNA      PAT      25-JAN-2001
LOCUS      AX070167
DEFINITION      Sequence 639 from Patent WO0102568.
ACCESSION      AX070167
VERSION      AX070167.1 GI:12579952
KEYWORDS
SOURCE
ORGANISM      human.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
AUTHORS      Williams,L.T., Escobedo,J., Inis,M.A., Garcia,P.D., Klinger,J.,
Kassam,A., Reinhard,C., Randazzo,F., Kennedy,G.C., Pot,D.,
Lamson,G., Drmanac,R., Crkenjakov,R., Drmanac,S., Dickson,M.,
Labat,I., Leshkowitz,D., Kita,D., Garcia,V. and Strache-Crain,B.
```

TITLE Human genes and gene expression products
JOURNAL Patent: WO 0102568-A 639 11-JAN-2001;
CHIRON CORPORATION (US); HYSEQ, INC. (US)
FEATURES
source 1. .685
/organism="Homo sapiens"
/db_xref="taxon:9606"
misc_feature 1. .685
/note="n = A,T,C or G"
BASE COUNT 238 a 127 c 125 g 187 t 8 others
ORIGIN

Query Match 18.9%; Score 513.8; DB 6; Length 685;
Best Local Similarity 94.9%; Pred. No. 1.6e-87;
Matches 560; Conservative 0; Mismatches 24; Indels 6; Gaps 3;

QY 558 tccaggggagagatccagatcaaaatgaaaaacagacacatctcgaatctcgaata 617
DB 1 tccaggggagagatccagatcaaaatgaaaaacagacacatctcgaatctcgaata 60
QY 618 ctgataacagagccagaaaaatccaaatgaaagccacttggggaagaaatcttaccctg 677
DB 61 ctgataacagagccagaaaaatccaaatgaaagccacttggggaagaaatcttaccctg 120
QY 678 acttacccttgcgcctcctcctcgaagaaactcaaaagagacatlaagacatctggagagc 737
DB 121 acttacccttgcgcctcctcctcgaagaaactcaaaagagacatlaagacatctggagagc 180
QY 738 gaggtaagaatcttcgagcaagatatacgtatccttattcctaataagaagaagccta 797
DB 181 gaggtaagaatcttcgagcaagatatacgtatccttattcctaataagaagaagccta 240
QY 798 aatttcacaaaccttgggtgcgaattctccctgtaccaaagtcacgaatctgcacatactg 857
DB 241 aatttcacaaaccttgggtgcgaattctccctgtaccaaagtcacgaatctgcacatactg 300
QY 858 cagaacacacatccatccacagccatgagtgagtgatcatttaagtcacagacagcag 917
DB 301 cagaacacacatccatccacagccatgagtgagtgatcatttaagtcacagacagcag 360
QY 918 tctgttaagcagaggaatattatattgaaagaaagcctaagaagacatgatttattc 977
DB 361 tctgttaagcagaggaatattatattgaaagaaagcctaagaagacatgatttattc 420
QY 978 ctccaatagtatattcacaatgcctgtgcacatggaggaataaatcttcacatgacg 1037
DB 421 ctccaatagtatattcacaatgcctgtgcacatggaggaataaatcttcacatgacg 480
QY 1038 acattgatacactatgacaaagaaagaaagagtgattattactcaagaatcaagta 1097
DB 481 acattgatacactatgacaaagaaagaaagagtgattattactcaagaatcaagta 536
QY 1098 ctccagtaagagatggggcaaaagagtgtagtgatgacaaaaaaca 1147
DB 537 ctccagtaagagatggggcaaaagagtgtagtgatgacaaaaaaca 584

RESULT 15
AC074175 177899 bp DNA 15-JUL-2000
LOCUS Mus musculus chromosome 5 clone RP23-348B11 strain C57BL6/J,
DEFINITION WORKING DRAFT SEQUENCE, 15 unordered pieces.
AC074175.1 GI:9111286
VERSION AC074175.1 GI:9111286
KEYWORDS HTG, HTGS, PHASE1, HTGS, DRAFT.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Granata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 177899)
Beckstrom-Sternberg, S.M., Benjamin, B., Blakeley, R.W.,
Bouffard, G.G., Dietrich, N.L., Eagle, W.O., Gupta, D., Ho, S.-L.,

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL

COMMENT

Huang, M.C., Idol, J., Lee-Lin, S.-Q., Maduro, O.L., Maduro, V.B.,
Mastrián, S.D., McCloskey, J.C., Ojeda, M.A., Pearson, R.,
Stantir, P.S., Summers, T.J., Thomas, J.W., Thomas, P.J.,
Tiongson, E.E., Touchman, J.W., Tran, J.T., Vogt, J.L., Walker, M.A.,
Wetherby, K.D., and Green, E.D.
NISC Mouse Sequencing Initiative
Unpublished
2 (bases 1 to 177899)
Direct Submission
Genome Center
Submitted (15-JUL-2000) NIH Intramural Sequencing Center, 8717
GroveMont Circle, Gaithersburg, MD 20877, USA

Center: NIH Intramural Sequencing Center
Center code: NISC
Web site: http://www.nisc.nih.gov
Contact: nisc_mouse@nigrl.nih.gov

Project Information
Center project name: VC
Center clone name: 348B11

Summary Statistics
Sequencing Vector: Plasmid
Chemistry: Dye-terminator Big Dye 100% of reads
Assembly program: Phrap version 990119
Consensus quality: 17064 bases at least Q40
Consensus quality: 172951 bases at least Q30
Consensus quality: 174122 bases at least Q20
Insert size: 16700; agarose-gel
Insert size: 20200; agarose-gel
Insert size: 176499; sum-of-ctrls
Quality coverage: 4.80x in Q20 bases; pulse-field-gel
Quality coverage: 5.16x in Q20 bases; sum-of-ctrls

NOTE: This is a 'working draft' sequence. It currently
consists of 15 contigs. The true order of the pieces
is not known and their order in this sequence record is
arbitrary. Gaps between the contigs are represented as
runs of N, but the exact sizes of the gaps are unknown.
This record will be updated with the finished sequence
as soon as it is available and the accession number will
be preserved.

1 2463: contig of 2463 bp in length
2464 2563: gap of unknown length
2564 4767: contig of 2204 bp in length
4768 4867: gap of unknown length
4868 8397: contig of 3530 bp in length
8398 8497: gap of unknown length
8498 13167: contig of 4670 bp in length
13168 13267: gap of unknown length
13268 17790: contig of 4523 bp in length
17791 17890: gap of 5666 bp in length
17891 23556: contig of 5666 bp in length
23557 23556: gap of unknown length
23557 32999: contig of 9343 bp in length
32999 33000: gap of unknown length
33000 44567: contig of 11468 bp in length
44567 44567: gap of unknown length
44568 56459: contig of 11792 bp in length
56459 56459: gap of unknown length
56460 72129: contig of 15570 bp in length
72130 72129: gap of unknown length
72130 87873: contig of 15644 bp in length
87874 87873: gap of unknown length
87874 108836: contig of 20863 bp in length
108837 108836: gap of unknown length
108837 130636: contig of 21700 bp in length
130637 130736: gap of unknown length
130737 151003: contig of 20267 bp in length
151004 151103: gap of unknown length
151104 177899: contig of 26796 bp in length.
Location/Qualifiers
1. .177899

FEATURES
source

	/organism="Mus musculus"
	/strain="C57BL6/J"
	/db_xref="taxon:10090"
	/chromosome="5"
	/clone="RP23-348B11"
	/clone_1lb="RPCI mouse BAC library 23"
	1. 2463
misc_feature	/note="assembly_fragment"
misc_feature	2564. .4767
	/note="assembly_fragment"
	clone_end:r7
	vector_side:left
misc_feature	4868. -8397
	/note="assembly_fragment"
misc_feature	8498. .13167
	/note="assembly_fragment"
misc_feature	13268. .17790
	/note="assembly_fragment"
misc_feature	17891. .23556
	/note="assembly_fragment"
	clone_end:sp6
	vector_side:right
misc_feature	23657. .32999
	/note="assembly_fragment"
misc_feature	33100. .44567
	/note="assembly_fragment"
misc_feature	44668. .56459
	/note="assembly_fragment"
misc_feature	56560. .72129
	/note="assembly_fragment"
misc_feature	72250. .87873
	/note="assembly_fragment"
misc_feature	87974. .108836
	/note="assembly_fragment"
misc_feature	108937. .130636
	/note="assembly_fragment"
misc_feature	130737. .151003
	/note="assembly_fragment"
misc_feature	151104. .177899
	/note="assembly_fragment"
BASE COUNT	53941 a 34927 c 35777 g 51822 t 1432 others
ORIGIN	
Query Match	15.6%; Score 423.6; DB 2; Length 177899;
Best Local Similarity	64.3%; Pred. No. 2.5e-70;
Matches 779; Conservative	0; Mismatches 394; Indels 39; Gaps 8;
OY 1502	aagaataaaatcacagtgcttgcattcccttcctcgtttctgcaagtctcctgaanaa 1561
Db 98960	AAAGAAATGAATGACTAGTGTGGGTCTTTCTTTCTTGTTTCGCAATGTCTGAAAAA 99039
OY 1562	ctgacaaaagaanaatgtgaattgccacaatatlttcacgaagaattgccagaagaatg 1621
Db 99040	CTGCACCGAAGAGAAGCCACTCTTGAGGCCCAATTCTCAGAAAGACGCTGGCGAAAGCA 99099
OY 1622	atacaacagtagaaggacagaattccctgtataaagagaccagaanaactgaaaaaagc 1681
Db 99100	GTCGGCACCTGCTCAAGCCGAACCTCTCATGTATGAAGAAACCAGAAACCTGTAAGAAAAC 99159
OY 1662	tccgttatattcagagcccatccccacccttcacaatgaattgagagggcttaatlaga 1741
Db 99160	ATGGGTTTGCTTCAAGAGCCCAACACCTATTCTTTCTTAGCTGAGATTGAAGAGATGTA----- 99214
OY 1742	aatagagtaataaagtccatgyltaagttaaacgcgtgaagatgacataagaacgaattta 1801
Db 99215	----CAGAAACAACATGAGCAATGTTCATGTGAAGAGCCCTGACCCAGAGCGAGTAAG 99270
OY 1802	cacgactaccttacataaaacaacaacagaaatgcatctcttgacatttcgcgaacacat 1861
Db 99271	CACACTACTCTTGGCTGTACAGCACACACCGGAA-----CACCAAGTGACCG 99315
OY 1862	taagtgaataaagacttaagaagaactaaggtgatgacatcataaattglaacatalacagcat 1921

[illegible]

Search completed: December 27, 2001, 20:56:56
Job time: 14571 sec

Fri Dec 28 08:22:47 2001

us-09-830-647-4.rge

Page 24

